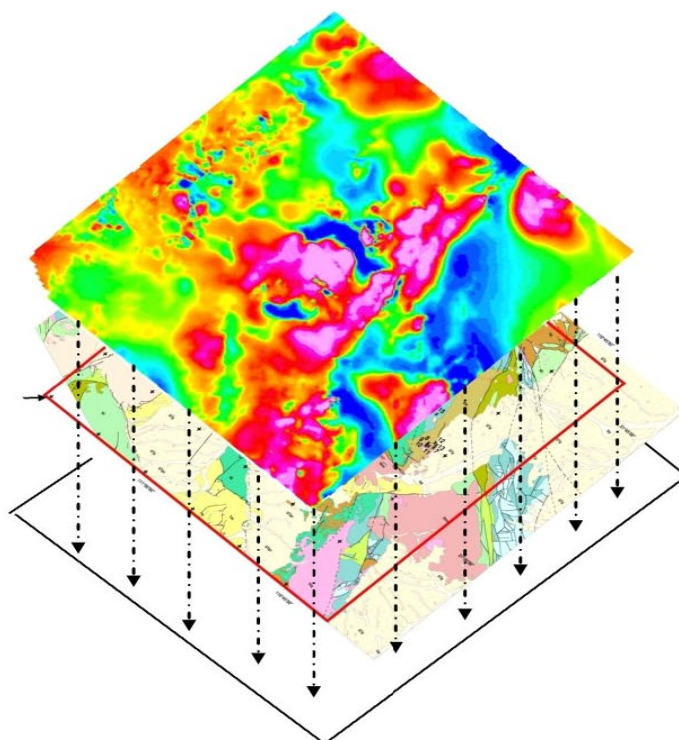




Texture-Based Automated Lithological Classification Using Aeromagnetic Anomaly Images

By Vivek Shankar



A thesis submitted to the faculty of the Department of Electrical and Computer Engineering, in partial fulfillment of the requirements for the degree of Master of Science, Graduate College, The University of Arizona, 2004

Open-File Report 2009-1206

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
KEN SALAZAR, Secretary

U.S. Geological Survey
Marcia K. McNutt, Director

U.S. Geological Survey, Reston, Virginia 2009

For product and ordering information:
World Wide Web: <http://www.usgs.gov/pubprod/>
Telephone: 1-888-ASK-USGS

For more information on the USGS—the Federal source for science about the Earth,
its natural and living resources, natural hazards, and the environment:
World Wide Web: <http://www.usgs.gov/>
Telephone: 1-888-ASK-USGS

Fiscal support for the work was provided from the USGS Mineral Resources Program
and a project grant from the National Park Service.

Suggested citation:
Shankar, Vivek, 2009, Texture-based automated lithological classification using aeromagnetic anomaly
images: U.S. Geological Survey Open-File Report 2009-1206 [<http://pubs.usgs.gov/of/2009/1206/>].

Any use of trade, product, or firm names is for descriptive purposes only and does not imply
endorsement by the U.S. Government.

Although this report is in the public domain, permission must be secured from the individual
copyright owners to reproduce any copyrighted material contained within this report.

Cover: Composite aeromagnetic image (from figure 5.5)

TEXTURE-BASED AUTOMATED LITHOLOGICAL CLASSIFICATION
USING AEROMAGNETIC ANOMALY IMAGES

By

Vivek Shankar

A Thesis Submitted to the Faculty of the

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

In Partial Fulfillment of the Requirements
For the Degree of

MASTER OF SCIENCE
WITH A MAJOR IN ELECTRICAL AND COMPUTER ENGINEERING

In the Graduate College

THE UNIVERSITY OF ARIZONA

2004

STATEMENT BY THE AUTHOR

This thesis has been submitted in partial fulfillment of requirement for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of the manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of the scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: _____

APPROVAL BY THESIS DIRECTORS

This thesis has been approved on the date shown below:

Jeffrey J. Rodriguez
Associate Professor of
Electrical and Computer Engineering

Date

Mark E. Gettings
Adjunct Professor of
Mining and Geological Engineering
Research Geophysicist, U.S. Geological Survey

Date

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisors Dr. Jeffrey J. Rodriguez and Dr. Mark E. Gettings, for their excellent guidance and support throughout the duration of this study. I would like to thank Dr. Bane Vasic for agreeing to be part of my committee and volunteering his valuable time.

I would also like to thank Douglas Hirschberg for helping me with the intricacies of the Arc GIS toolbox and Ali Tabesh for his time and infinite patience in answering my unending list of questions.

Finally, I would like to thank my family and friends for their constant encouragement.

TABLE OF CONTENTS

	LIST OF FIGURES.....	6
	LIST OF TABLES	7
	ABSTRACT.....	8
1	INTRODUCTION.....	9
2	GEOPHYSICAL PROSPECTING, MAGNETISM AND AEROMAGNETIC ANOMALY SURVEYING.....	13
2.1	Mineral Resources	13
2.2	Geophysical Prospecting.....	14
2.3	Problem of Ambiguity.....	16
2.4	Rock Magnetism	17
2.4.1	Theory of Magnetism	17
2.4.2	Magnetism at Atomic Scale.....	19
2.4.3	Magnetic Minerals.....	20
2.5	Geomagnetism.....	21
2.6	Magnetic Anomalies and Aeromagnetic Anomaly Surveys.....	24
2.6.1	Magnetic Anomalies.....	24
2.6.2	Data Acquisition.....	26
2.6.3	Aeromagnetic Survey Specifications.....	28
2.6.4	Data Reduction.....	29
2.6.5	Data Presentation.....	30
2.6.6	Applications of Aeromagnetic Anomaly Surveys.....	32
3	TEXTURE ANALYSIS	33
3.1	Definition of Texture	33
3.2	Taxonomy for Texture	36
3.3	First-Order Statistics.....	38
3.4	Second-Order Statistics.....	39
3.5	Gettings Measures.....	43
3.6	Texture Spectrum Based Measures.....	46
3.7	Fractal Dimension Based Measures.....	51
4	STATISTICAL PATTERN CLASSIFICATION	54
4.1	Introduction.....	54
4.2	Classifiers.....	58
4.2.1	Bayesian Classifier.....	58
4.2.2	Bayes Decision Rule for Minimum Error Classification.....	59
4.2.3	Discriminant Functions.....	60
4.2.4	Gaussian Classifier.....	61
4.2.5	<i>k</i> -Nearest Neighbor Classifier.....	62
4.3	Feature Selection and Error Estimation.....	64
4.3.1	Feature Selection Criterion and Strategy.....	65
4.3.2	Feature Extraction.....	67
4.3.3	Error Estimation.....	68
5	DATA AND SOFTWARE	69
5.1	Data.....	69
5.1.1	Aeromagnetic Anomaly Data.....	69

5.1.2	Geologic Data.....	72
5.1.3	Composite Aeromagnetic Image.....	74
5.1.4	Brodatz Images.....	76
5.2	Software.....	76
5.2.1	Geophysical Package.....	76
5.2.2	Geological Package.....	76
5.2.3	Feature Measurement Code.....	77
5.2.4	Pattern Classification Toolbox and Data Visualization.....	78
6	EXPERIMENTS AND RESULTS	81
6.1	Experiments Using Survey and Gridded Data.....	81
6.2	Classification Using Complete Feature Vector.....	83
6.2.1	Class Conditional Feature Histograms.....	83
6.2.2	Accuracy Results for Complete Feature Vector.....	87
6.3	Using Feature Selection Strategies.....	87
6.3.1	Issues with Experimental Setup.....	88
6.3.2	Experiments Using Non-Overlapping Window.....	89
6.4	Experiments with Spatially Uncorrelated Sets.....	89
6.4.1	Experiments with Two Spatially Uncorrelated Sets.....	89
6.4.2	Experiments with Three Spatially Uncorrelated Sets.....	92
6.5	Data Visualization.....	93
6.5.1	Scatter Plots, Convex Hulls and d^2 -display.....	93
6.5.2	Experiments without Class 1050.....	97
6.6	Grouping of Units.....	98
6.6.1	Groups.....	98
6.6.2	Experiments Conducted Post-Grouping	98
6.6.3	Experiments Using Pure Training Samples.....	99
6.7	Effect of Varying Window Size.....	100
6.8	Effect of Varying Quantization Levels.....	101
6.9	Comparative Performance of Features.....	102
6.10	Discussion of Results.....	105
6.10.1	Experimental Setup.....	105
6.10.2	Classification Accuracy.....	105
6.10.3	Comparison with Previous Work.....	106
6.10.4	Comparison with Similar Quantitative Studies.....	107
7	CONCLUSION AND FUTURE WORK	108
	APPENDIX A: Creating a Composite Aeromagnetic Image	110
	APPENDIX B: Lithological Units in Patagonia	120
	APPENDIX C: Feature Vector Elements	123
	APPENDIX D: List Of Features Chosen	124
	REFERENCES	125

LIST OF FIGURES

2.1	Magnetic fields produced by different sources.....	18
2.2	Arrangement of atomic dipoles in different types of magnetic material.....	20
2.3	Simplistic model of the geomagnetic field.....	22
2.4	Geomagnetic field elements.....	23
2.5	Magnetic anomalies and their measurement.....	25
2.6	Fixed wing magnetometer.....	27
2.7	Traverse lines of airborne and ground magnetic surveys.....	28
2.8	Pseudo-colored image rendition of aeromagnetic anomaly survey data.....	31
3.1	Taxonomy of texture.....	36
3.2	Taxonomy of texture descriptors.....	36
3.3	Alternate classification of texture measures.....	37
3.4a	Example image.....	41
3.4b	For example image, using $d = 1$	41
3.5	Three functions with identical mean and standard deviation.....	44
3.6	Aeromagnetic anomaly flight line data.....	45
3.7	Transforming a neighborhood into a texture unit number.....	48
3.8	Calculating the Hurst coefficient.....	52
4.1	Model for automated classification.....	55
4.2	Choosing the best feature subset based on peaking accuracy.....	66
5.1	Aeromagnetic survey location.....	69
5.2	Pseudo-colored aeromagnetic anomaly image.....	71
5.3	Gray-scale version of the aeromagnetic anomaly image.....	72
5.4	Colored geological map of the Patagonia area.....	73
5.5	Composite aeromagnetic image.....	75
6.1	Class conditional feature histograms for class 1050.....	84
6.2	Class conditional feature histograms for class 1056.....	85
6.3	Class conditional feature histograms for class 1082.....	86
6.4	Scatter plot and convex hull for set 1.....	94
6.5	Scatter plot and convex hull for set 2.....	95
6.6	d^2 -display for class 1050 vs. rest of the data.....	96
6.7	Reasons why EL and EL/NPT perform better than EL alone.....	104

LIST OF TABLES

2.1	Geophysical surveying methods.....	15
2.2	Geophysical surveying applications.....	16
5.1	Fields in the aeromagnetic anomaly database.....	70
5.2	Computational complexity for feature measurement code.....	78
5.3	Computational complexity for TOOLDIAG experiments.....	79
6.1	Summary of lithological units.....	82
6.2	Classification accuracy for survey intensity values.....	82
6.3	Classification accuracy for gridded intensity values.....	83
6.4	Classification accuracy for 49 features.....	87
6.5	Classification accuracy after feature selection.....	88
6.6	Summary of lithological units for the non-overlapping window experiment	90
6.7	Classification accuracy after feature selection for non-overlapping window	90
6.8	Summary of lithological units for 2 ways split.....	91
6.9	Classification accuracy for set 2.....	91
6.10	Summary of lithological units for 3 ways split.....	92
6.11	Classification accuracy for set 3.....	93
6.12	Summary of lithological units for 3 ways split without class 1050.....	97
6.13	Classification accuracy for set 3 without 1050.....	97
6.14	Grouping of units.....	98
6.15	Summary of lithological units after grouping.....	99
6.16	Classification accuracy after grouping.....	99
6.17	Classification accuracy for pure training samples.....	100
6.18	Effect of changing window size.....	100
6.19	Effect of the number of quantization levels.....	101
6.20	Best performing features.....	102

ABSTRACT

Aeromagnetic anomaly images are geophysical prospecting tools frequently used in the exploration of metalliferous minerals and hydrocarbons. The amplitude and texture content of these images provide a wealth of information to geophysicists who attempt to delineate the nature of the Earth's upper crust. These images prove to be extremely useful in remote areas and locations where the minerals of interest are concealed by basin fill. Typically, geophysicists compile a suite of aeromagnetic anomaly images, derived from amplitude and texture measurement operations, in order to obtain a qualitative interpretation of the lithological (rock) structure. Texture measures have proven to be especially capable of capturing the magnetic anomaly signature of unique lithological units. We performed a quantitative study to explore the possibility of using texture measures as input to a machine vision system in order to achieve automated classification of lithological units. This work demonstrated a significant improvement in classification accuracy over random guessing based on *a priori* probabilities. Additionally, a quantitative comparison between the performances of five classes of texture measures in their ability to discriminate lithological units was achieved.

Chapter 1

INTRODUCTION

Geophysical prospecting methods have become increasingly important tools for mineral and fossil fuel exploration. These methods involve using a physical property of the Earth, such as density or magnetization, to extract structural and stratigraphic information of a region [1]. Mineral resources of interest often produce small magnetic fields that distort the main magnetic field of the Earth. A geophysical prospecting technique that involves the aerial capture of these perturbations in the Earth's magnetic field over a region of interest is *aeromagnetic anomaly surveying*. The work product of such a survey is a non-uniformly sampled two-dimensional array of aeromagnetic anomaly intensities. The geophysicist typically renders this data as a pseudo-colored image to overlay onto the geologic map of the region for use in subsequent interpretation of lithological structure. Several amplitude-based transformations, inspired by standard signal and image processing techniques, are employed to highlight useful information in the aeromagnetic data. Discussions on various linear and non-linear transformations, including wave number, gradient and automatic gain control based methods, can be found in publications by Dentith [2] and Milligan [3].

When interpreting aeromagnetic anomaly and derivative images, geophysicists delineate areas with similar amplitude and texture characteristics, and these areas usually correspond to particular lithological units. Dentith [4] observed that although texture was an important consideration while interpreting aeromagnetic data, the commonly used transformations were designed to enhance the aeromagnetic anomaly images based on

variations in amplitude. In order to better utilize texture information in his study on gold mineralization in Western Australia [4], he used four methods of quantifying texture that are widely used in remote sensing and biomedical applications of image processing. He generated derivative anomaly images using first-order statistics, gray-level co-occurrence matrices (GLCM), texture spectrum and fractal dimension based transforms. On studying the derivative images, Dentith found that when mineralization occurs in strongly magnetic rocks, amplitude-based transforms are successful in highlighting areas and structures that are likely to be mineralized. However, when mineralization occurs within rocks which lack magnetization contrasts, there is little amplitude variation and in such areas texture-based transforms prove to be more useful. Further, Dentith's subjective, qualitative analysis led to the conclusion that texture transforms based on GLCM and fractal dimension generate the most effective derivative images to highlight useful lithological information.

To aid in his study of the Patagonia region of southern Arizona, Gettings [5] devised two wavelet-based texture measures, namely Euclidean length and number of peaks and troughs, to generate derivative aeromagnetic anomaly images. Correlating these derived images with the geological map of the region revealed that the presence of several lithological structures was corroborated accurately by trends in the derived images. Taking his analysis a step beyond what Dentith attempted, Gettings employed a ranking system [5] to generate derivative images based on a combination of the two aforementioned wavelet-based measures. He reported that these derived images provided useful lithological information not delineated when the measures were used by themselves. Gettings concluded that texture analysis of aeromagnetic anomaly images

showed great promise and that the methodology used to combine measures required more in-depth study.

Dentith and Gettings' studies demonstrated that textural characteristics of aeromagnetic data could provide insightful lithological information, especially in areas lacking in magnetization contrast. However, Dentith studied each of the derivative images in isolation, while Gettings' attempt at combining measures, using his ranking system, to generate derivative images can only be viewed as a first step towards taking a multi-measure (or multi-dimensional) approach to the problem. Further, both studies were subjective and qualitative in nature, and did not provide any quantitative basis to justify the use of texture measures.

As a logical follow-up to Gettings' work in the Patagonia region of southern Arizona [5], we framed a classical pattern recognition problem to determine whether the aeromagnetic anomaly data of the region could be used to quantitatively predict the lithological units. The aeromagnetic anomaly survey data was uniformly re-sampled to create a more convenient image. A digital geologic map of the surveyed region indicated the lithological units present. The information from both these sources formed a composite dataset, consisting of the anomaly reading and lithological unit at every pixel location in the image. The composite was used as the input to a feature measurement algorithm. Five sets of texture measures, totaling 49 features, were computed for each pixel location using a sliding window. These features, along with the known lithological unit labels, were then split into training, evaluation and test sets. A rigorous statistical analysis followed, involving experimentation with a combination of feature selection and classification schemes to determine the highest prediction accuracy that could be

achieved. Visualization tools were used to shed light on the manner in which the lithological units cluster in multi-dimensional texture space. An important byproduct of this work was a comparative study of the quantitative ability of different types of texture measures in discriminating lithological units. Finally, the statistical analysis yielded some important geological inferences about the region, which corroborates observations made by Gettings [5].

Chapter 2 of this report discusses the physics involved in geophysical prospecting and magnetism. Details are provided about the nature of magnetic anomalies and the methods used to acquire the data. Chapter 3 describes the methods of quantifying texture. The definitions of various texture measures used in this study are listed. Chapter 4 delves into statistical pattern recognition theory. The statistical tools used in this study, including feature selection algorithms, parametric and non-parametric classification schemes and visualization methods are detailed. Chapter 5 lays out the experimental setup, discusses the assumptions and limitations of the study and describes the data, tools and software used in this work. Chapter 6 provides a detailed look at the various experiments conducted, tables of accuracy numbers, data visualization graphs and other results. Chapter 7 includes the conclusions drawn and potential future work. The appendices include a user manual detailing step-by-step procedures to run various functions available in the software tools, as well as lists of the lithological units present in the Patagonia area and texture measures used in the study.

Chapter 2

GEOPHYSICAL PROSPECTING, MAGNETISM AND AEROMAGNETIC ANOMALY SURVEYING

2.1 Mineral Resources

Mining is the process of extracting useful minerals, including metals, non-metals and hydrocarbons, from the subsurface of the Earth and sea beds. In their natural form mineral resources exist as an aggregate of the useful minerals and gangue (associated rock of no economic value) [1]. *Reserves* are mineral resources that are known to exist with or without the assurance that they can be exploited economically [6] [7]. Appraisal of unknown resources for future consumption, a key part of several governments and corporations' mineral resource agenda, is referred to as *prospecting* or *exploration*. Traditionally, prospectors relied on direct observation of mineralization in the rock outcrops, sediments and soil. Direct methods involve excavating and drilling boreholes. Although these methods would unambiguously delineate the nature of the Earth's crust, they pose logistic problems. Excavation would be cost prohibitive considering the surface area in question and the difficulty in accessing several remote locations. Boreholes would provide information only at discrete locations. Indirect approaches, such as a combination of geologic, geochemical and *geophysical* prospecting tools, have been increasingly employed. Although prone to ambiguities of interpretation, they provide a rapid and cost-effective means of deriving spatially distributed information on subsurface mineral resources. Prospecting is not a substitute for drilling; rather, when properly applied it can

optimize the exploration program by maximizing the rate of ground coverage while minimizing the drilling requirement [1].

2.2 Geophysical Prospecting

Geophysical prospecting involves the application of the principles of physics to the study of the Earth [1]. Measurements are taken at or near the surface of the Earth and analysis reveal how the physical properties of the Earth's interior vary vertically and laterally. Studies could cover the Earth in its entirety (global geophysics) or a local region of the upper crust. In several geophysical surveying methods, it is the local variations in the measured parameter, relative to some normal background value, that is of primary interest. Such variation is attributable to a localized subsurface zone of distinctive physical property and possible geological importance. A local variation of this type is known as a *geophysical anomaly*.

Prospecting methods are classified according to the fields of the Earth that they measure. Table 2.1 lists the popular geophysical prospecting methods. The methods can also be divided into natural and artificial field methods. Natural field methods use the naturally occurring fields, such as gravitational, magnetic, electrical and electromagnetic fields of the Earth, to help predict concealed geological features. Artificial source methods involve generating local seismic, electrical or electromagnetic fields that may be used analogously to natural fields. An example of an artificial method is seismic surveying. In this method, seismic waves are generated whose propagation velocities and transmission paths through the subsurface are mapped to provide information on the distribution of geological boundaries at different depths.

Method	Measured Parameter	Operative Physical Property
Seismic	Travel times of reflected/refracted seismic waves	Density and elastic moduli, determines the propagation velocity of seismic waves
Gravity	Spatial variations in the strength of the gravitational field of the Earth	Density
Magnetic	Spatial variations in the strength of the geomagnetic field	Magnetic susceptibility and remnant magnetization
Electrical - Resistivity, induced polarization or self potential	Earth resistance, polarization voltage or electrical potentials	Electrical conductivity, capacitance
Electromagnetic	Response to electromagnetic radiations	Electrical conductivity and inductance
Radar	Travel times of reflected radar pulses	Dielectric constants

Table 2.1: Geophysical surveying methods [1]

Geophysical surveys can be ground-based, airborne or ship-borne depending on the survey area. Ground geophysical surveys are conducted by transporting mobile instruments across the region of interest and taking measurements at predetermined locations. These types of surveys may be limited by accessibility issues and may be restricted to areas where public roads are available [8]. Airborne geophysical surveys do not suffer from this limitation. A probe is attached to an aircraft and the onboard instrumentation takes care of acquiring readings at regular intervals. Reconnaissance

surveys are usually carried out from the air because of the high speed of operation and the extended area that can be covered.

The type of physical property to which a prospecting method responds determines its applications. Table 2.2 lists surveying techniques used for some important prospecting applications. In our work, our primary interest was geophysical surveying based on airborne sensing of anomalies in the natural magnetic field.

Application	Appropriate Survey Method
Exploration of fossil fuel (oil, gas, coal)	S, G, M, EM
Exploration of metalliferous mineral deposits	M, EM, E, SP, IP, R
Exploration of bulk mineral deposits	S, E, G
Exploration for underground water	E, S, G, Rd
Engineering/construction site investigation	E, S, Rd, G, M
G=gravity, M=magnetic, S=seismic, E=electrical resistivity, SP=self potential, IP=induced polarization, EM=electromagnetic, R=radiometric, Rd=ground-penetrating radar.	

Table 2.2: Geophysical surveying applications [1]

2.3 Problem of Ambiguity

The Earth's internal structure is not precisely known. Hence, the source of any particular geophysical measurement taken at the Earth's surface cannot be predicted uniquely. In other words, the effect is known but the cause is not uniquely understood. Such a problem, in which the goal is to deduce a model for the unknown source, is

referred to as an *inverse problem*. Inverse problems suffer from an inherent ambiguity or non-uniqueness in conclusions that can be drawn from the data. This is primarily because different geological configurations could result in the same observed measurements. Additionally, finite precision information systems, calibration issues and other experimental inaccuracies add a degree of indeterminacy to the problem. As a unique solution is not usually available, the typical approach is either to determine properties of the subsurface that all possible solutions share or to introduce assumptions to restrict the number of admissible solutions. In spite of these inherent problems, geophysical surveying is an invaluable tool for investigation and plays a key role in exploration programs.

Since the data used in this study was acquired by using a geophysical prospecting method based on magnetism, namely aeromagnetic anomaly surveying, the rest of this chapter is devoted to the intricacies of magnetic fields and rock magnetism, followed by a discussion of the instrumentation and processes involved in conducting an aeromagnetic anomaly survey.

2.4 Rock Magnetism

2.4.1 Theory of Magnetism

Magnetism is a phenomenon associated with the motion of charged particles. Fundamental to magnetism is the *magnetic field* and its effects on matter, namely deflection of moving charges and torque on other magnetic materials. As depicted in Figure 2.1, a magnetic field is a region in the neighborhood of a magnet, electric current or changing electric field, in which magnetic forces are observable. The field around a

permanent magnet or a steady electric current is referred to as a *magnetostatic field* [1].

A fluctuating direct current or an alternating current results in a magnetic field that constantly changes in magnitude and direction.

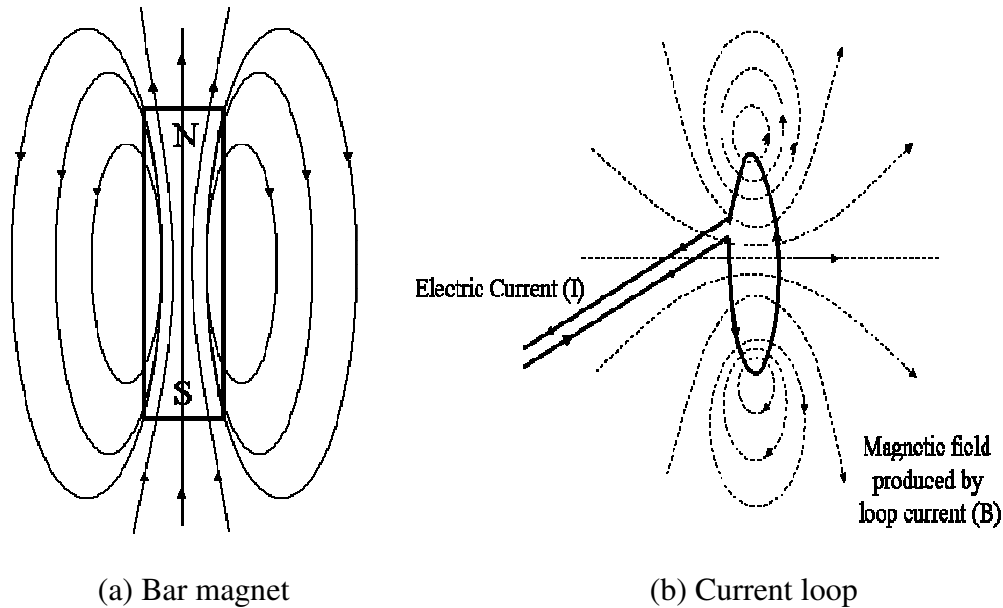


Figure 2.1: Magnetic fields produced by different sources [10]

Magnetic fields are represented by continuous lines of force that originate at the north magnetic pole (positive pole) and terminate at the south magnetic pole (negative pole). The density of the lines indicates the magnitude of the magnetic field. At the poles the density is high and the magnetic field is strong. Farther away, the lines fan out and the magnetic field is weak. Magnetic fields are usually represented using one of two vector quantities described next. *Magnetic field intensity* (also known as *magnetic field strength* or *magnetizing field*) H is the field produced by the flow of charged particles. *Magnetic flux density* or *magnetic induction* B is the total magnetic field including the field generated due to the magnetic properties of the material present in the magnetizing field.

In this study we ignore the directional property of the magnetic field and only use field strength in the analysis. The strength or magnitude of B is expressed in units of Tesla (T). Since Tesla is too large a unit in which to express the small magnetic fields caused by lithological units, nanotesla (nT) is commonly employed.

2.4.2 Magnetism at Atomic Scale

An investigation into the cause of the magnetism phenomenon takes us to the atomic level. All substances are magnetic at an atomic scale [1]. Each atom acts as a dipole due to both the spin of its electrons and the orbital path of the electrons around the nucleus. Quantum theory allows two electrons to exist in the same orbital provided they have spins in opposite directions. Two such electrons are called *paired electrons* and their spin magnetic moments cancel. In *diamagnetic* materials, like carbon and silicon, all electron orbitals are full and no unpaired electrons exist. Hence the *susceptibility* (measure of the degree of magnetization in response to a magnetic field) displayed by these materials is negative. However, when unpaired electrons exist in a material, there is a magnetic dipole associated with their spins. Figure 2.2 depicts the dipole arrangements in such materials. *Paramagnetic* materials, like magnesium and lithium, have incomplete electron orbitals. This causes an induced field in the same direction as the applied field. The susceptibility of these materials is positive but weak because the arrangement of the dipoles is not ordered. In *ferromagnetic* materials dipoles associated with the spins of unpaired electrons are parallel. This results in strong spontaneous magnetization which can exist even in the absence of external field. This type of material, like iron, cobalt and nickel, rarely occurs in its natural form. *Anti-ferromagnetic* materials, like haematite (ferric oxide), have anti-parallel dipole arrangements with equal number of dipoles in

each direction. Magnetic fields of these dipoles are self-canceling, resulting in a zero net induced magnetic field. However, defects in crystal lattice structure may give rise to a small net magnetization, called *parasitic anti-ferromagnetism*. Dipoles in *ferrimagnetic* materials, like magnetite (ferrous ferric oxide), arrange themselves in anti-parallel with the number of dipoles in each direction being unequal. This causes strong spontaneous magnetization and positive susceptibility.

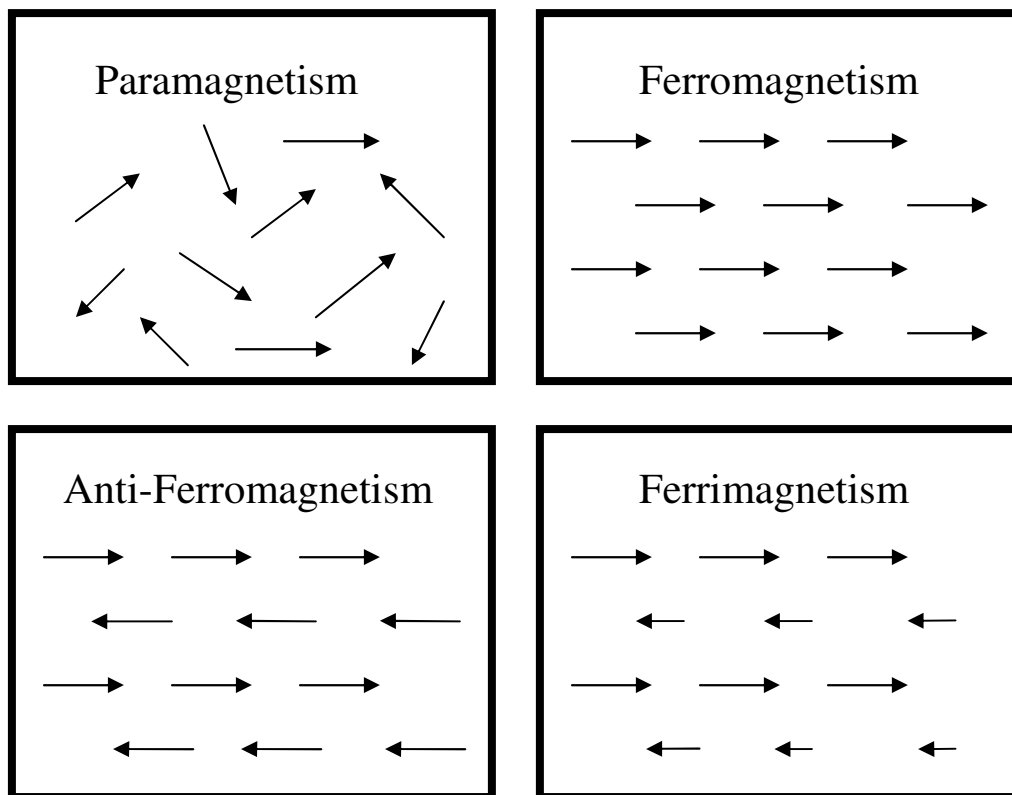


Figure 2.2: Arrangement of atomic dipoles in different types of magnetic material [7]

2.4.3 Magnetic Minerals

Common rock-forming minerals exhibit a very low magnetic susceptibility and owe their magnetic nature to the typically small proportion of magnetic minerals they

contain. The most common magnetic mineral is magnetite [1]. Although the size, shape and dispersion of the magnetite grains within a rock affect its magnetic character, it is reasonable to classify the magnetic behavior of rocks according to their overall magnetite content. Igneous rocks are usually highly magnetic due to their relatively high magnetite content. Sedimentary rocks are effectively non-magnetic unless they contain a significant amount of magnetite in their heavy mineral fraction. Typically, in locations where magnetic anomalies are observed over sediment-covered areas, the anomalies are generally caused by an underlying igneous or metamorphic basement. Metamorphic rocks display variable magnetic behavior that is bound by the aforementioned extremes.

The presence of a variety of all three kinds of rocks in the Patagonia region makes it a good choice for our study. More details on the geologic setting of Patagonia can be found in Gettings' report [5].

2.5 Geomagnetism

As depicted in Figure 2.3, the Earth has a substantial magnetic field often referred to as the *geomagnetic* or normal magnetic field. Magnetic field lines radiate between Earth's north and south magnetic poles just as they do between the poles of a bar magnet. However, geomagnetism cannot result from permanent magnetism in the Earth's deep interior. The required dipolar magnetic moments are far greater than is considered realistic, and the prevailing high temperatures are far in excess of the Curie temperature of any known magnetic material. The geomagnetic field is believed to be produced by several sources that are present above and below the surface [1] [10].

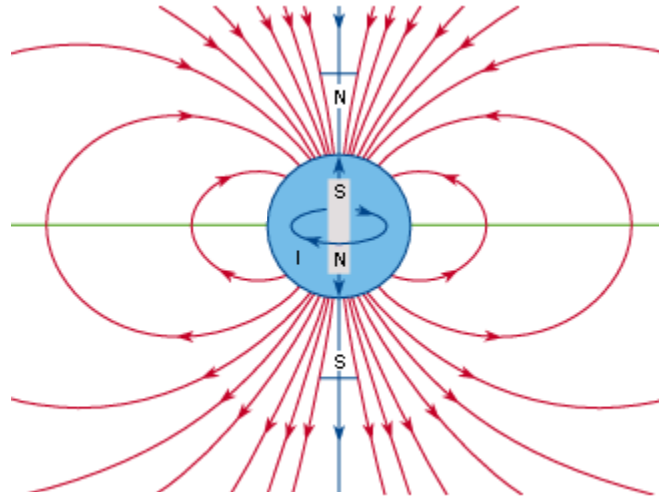


Figure 2.3: Simplistic model of the geomagnetic field [11]

The list of potential sources includes a geomagnetic dynamo, crustal magnetization, ionospheric dynamo, ring current, magnetopause current, tail current, field-aligned currents, and auroral electro-jets [11]. The main source of geomagnetic field is attributed to a dynamo action produced by the circulation of charged particles in coupled convective cells within the outer, fluid part of the Earth's core. The exchange of dominance between such cells is believed to produce the *periodic changes in the polarity* of the geomagnetic field revealed by the paleomagnetic studies. The circulation patterns within the core are not fixed and change slowly with time. As we move farther above the Earth's surface, the effects of other sources become stronger than that of the geomagnetic dynamo. The geomagnetic field is geometrically more complex than the gravity field of the Earth and exhibits irregular variation in both orientation and magnitude with latitude, longitude and time. The main field is perturbed by seasonal and solar cycle changes, solar and lunar tidal effects and the flow of charged particles within the ionosphere towards the magnetic poles.

At any point on the Earth's surface, a freely suspended magnetic needle assumes a position in space in the direction of the ambient geomagnetic field. This will generally be at an angle to both the vertical and geographic north. The *total field vector* F , as shown in Figure 2.4, has a vertical component Z and a horizontal component H in the direction of the magnetic north. The dip of F is the *inclination* L of the field, and the horizontal angle between geographic and magnetic north is the *declination* D . F varies in strength from about 25000 nT in the equatorial regions to about 70000 nT at the poles.

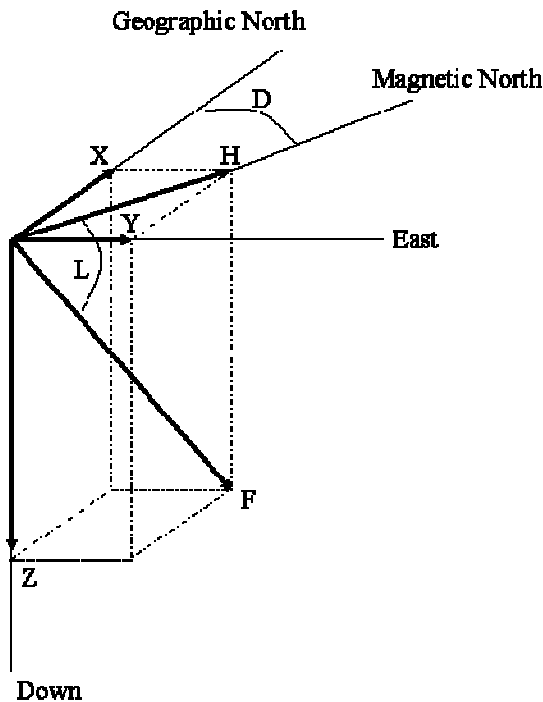


Figure 2.4: Geomagnetic field elements [10]

About 90% of the Earth's field can be represented by the field of a theoretical magnetic dipole at the center of the Earth inclined at about 11.5° to the axis of rotation. The magnetic moment of this fictitious *geocentric dipole* can be calculated from the

observed field. If this dipole field is subtracted from the observed magnetic field, the residual field can then be approximated by the effects of a second smaller, dipole. The process can be continued by fitting dipoles of ever decreasing moment until the observed geomagnetic field is simulated to any required degree of accuracy. The effects of each fictitious dipole contribute a function known as a harmonic, and the technique of successive approximations of the observed field is known as spherical harmonic analysis. This method is used to compute the *International Geomagnetic Reference Field* (IGRF) [1], which defines the theoretical undisturbed magnetic field at any point on the Earth's surface. Magnetic anomalies caused by rocks are localized effects superimposed on the geomagnetic field. Consequently, knowledge of the behavior of the geomagnetic field is necessary in interpretation of the magnetic data.

2.6 Magnetic Anomalies and Aeromagnetic Anomaly Surveys

2.6.1 Magnetic Anomalies

Geologic structures, including dykes, faults, folded or truncated sills and lava flows, massive basic intrusions, metamorphic basement rocks and magnetite ore bodies produce magnetic fields that distort the main magnetic field of the Earth. These distortions are called "anomalies" and can be detected by measuring the magnetic field near the surface of the Earth. Magnetic anomalies range from a few tens of nanotesla to several thousand nanotesla. Figure 2.5 shows the superposition of the magnetic anomalies onto the geomagnetic field and the aerial measurement process. By processing and analyzing these measurements, geophysicists can learn about geologic structures, even though the structures may be concealed entirely below the Earth's surface.

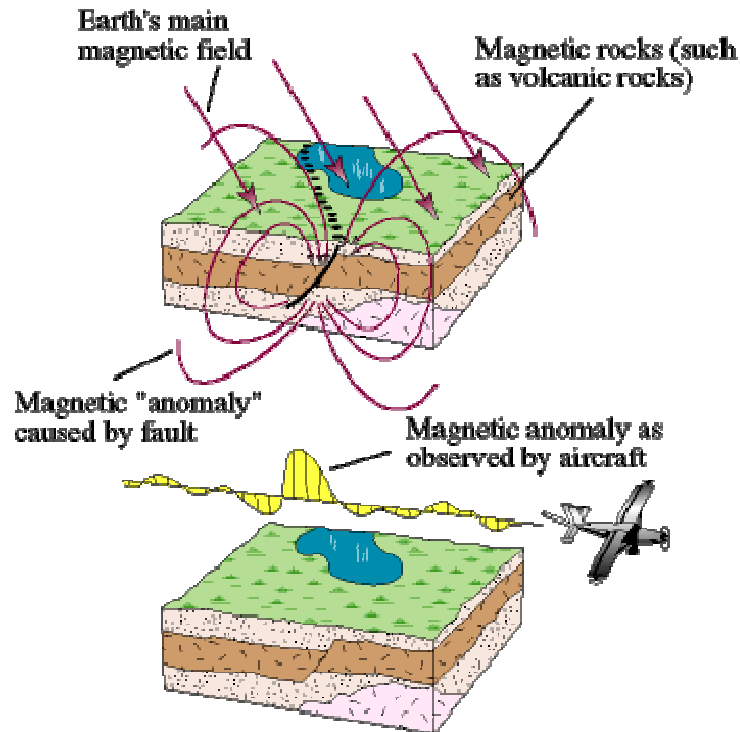


Figure 2.5: Magnetic anomalies and their measurement [12]

The magnetic field, measured at any point near the Earth's surface, is the vector sum of the following:

1. The Earth's geomagnetic field due to the dynamo action
2. An induced field caused by magnetic induction in magnetically susceptible materials polarized by the main field
3. A field caused by remnant magnetization of materials
4. Less significant fields caused solar, atmospheric and cultural influences.

The induced and remnant fields are of particular interest because the magnitudes of these fields are directly related to the magnetic susceptibility, spatial distribution and concentration of the local crustal materials. Once the main field and the minor source

effects are removed from the observed magnetic field data via various data reduction and processing methods, the processed data serves as an indicator of the spatial distribution and concentration of the magnetically significant minerals.

2.6.2 Data Acquisition

In some applications, the magnetic field data is acquired close to ground level either by a person carrying a magnetometer or with a magnetometer mounted on a vehicle such as a quad motorcycle or a four-wheel drive. Such a survey is referred to as a *ground magnetic* survey. Ground magnetic surveys are usually conducted over relatively small areas or locations where the target is defined and a more detailed picture of the magnetic field needs to be created. However, the majority of magnetic surveys are carried out in the air. Such a survey is referred to as an *aeromagnetic anomaly* survey. Aeromagnetic surveys are rapid and cost effective, typically costing 40% less per line kilometer than a ground survey [1]. Large areas can be covered without the cost of sending a field party into the survey area, and data can be obtained in areas that are inaccessible to ground surveys.

The magnetic field is usually measured by mounting a total field magnetometer on a fixed wing aircraft or helicopter. Fixed wing acquisition, as shown in Figure 2.6 is preferred due to the lower cost, though helicopters are necessary where the terrain is rugged. The sensor is well shielded to protect the instrument from the magnetic effects of the aircraft. An on-board recording system is present to store the readings. The most common instrument in use today is the cesium vapor magnetometer which has a magnetic field measurement precision of better than ± 0.01 nT.



Figure 2.6: Fixed wing magnetometer [14] (with a sensor in the stinger behind the tail)

Observations are made at regular intervals, between 1 and 15 meters, along a series of traverse lines, or *flight lines* of constant spacing as shown in Figure 2.7. Observations are similarly made along tie lines oriented perpendicular to the flight lines. Tie lines are necessary to assist in the removal of temporal variations in the main field. Tie lines are usually spaced ten times further apart than flight lines. While data is being collected along the flight and tie lines, a base station magnetometer also measures the magnetic field at a stationary point. This data serves as an estimate of the temporal variation of the main field, which is subtracted from the survey data. The base station magnetometer is also used to identify magnetic storm events, when the magnetic field is varying rapidly due to disturbances in the ionosphere or magnetosphere. On such occasions data acquisition is suspended as the estimate of the temporal variation is less accurate at a distance from the base station.

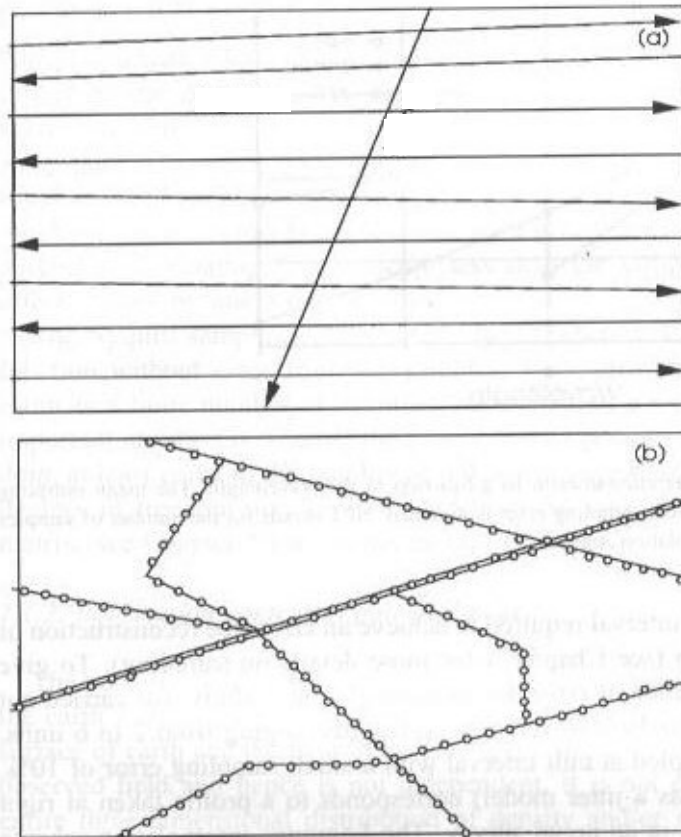


Figure 2.7: Traverse lines of (a) airborne and (b) ground magnetic surveys [8]

2.6.3 Aeromagnetic Survey Specifications

Aeromagnetic survey specifications are determined by consideration of several factors [13]. We now discuss a few of the more important specifications.

Survey Resolution – there is a tradeoff between cost of the survey and the detail captured, which in turn determines the flight-line spacing. Smaller line spacing implies higher cost but also higher resolution. Flight line spacing ranges from 400 m in regional mapping programs down to 50 m for detailed mapping projects at prospect scale. Finer lines spacing, of around 20 m, is sometimes employed in special circumstances.

Survey Orientation – maximum information is extracted when survey lines are oriented perpendicular to the geological strike (structures of most interest).

Survey Height – the main factor governing survey height (terrain clearance) is safety, which depends on the ruggedness of the terrain and the climbing capability of the aircraft. Except for the safety factor, surveys are carried out at a constant, and lowest possible, terrain clearance. However, due to the difficulty of processing and interpreting data acquired at highly variable clearances, a flying clearance that the aircraft can comfortably manage is chosen.

2.6.4 Data Reduction

Data reduction is the series of steps taken to remove both signal and spurious noise from the data that are not related to the geology of Earth's crust [13]. This process prepares the dataset for interpretation by reducing the data to only contain signals relevant to the task. The main reduction steps are as follows:

Magnetic compensation – it involves removal of influences of the magnetic signature of the aircraft on the recorded data and is often done in real time onboard the aircraft.

Cultural influences – it involves the removal of spurious noise and spikes from the data caused by power lines, metallic structures, radio transmissions, fences, etc.

Diurnal corrections – it involves corrections for the temporal variation of the Earth's main field. This is achieved by subtracting the time-synchronized signal, recorded at a stationary base magnetometer, from the survey data. This procedure relies on the assumption that the temporal variation of the main field is the same at the base station and in the survey area. Best results are obtained if the base station is close to the survey area.

IGRF removal – geomagnetic reference field removal removes the strong influence of the Earth's main field on the survey data. This is done because the main field is dominantly influenced by dynamo action in the core and not related to the geology of the (upper) crust. This is achieved by subtracting a model of the main field from the survey data.

Tie line leveling – additional data recorded on tie lines is used to further adjust the data. Data recorded at intersections (crossover points) of flight and tie lines should be equal.

Micro-leveling – it involves the removal of any errors remaining after all the aforementioned adjustments are applied. This step usually takes care of the very subtle errors caused by variations in terrain clearance or elevated diurnal activity.

2.6.5 Data Presentation

Magnetic anomaly data acquired along flight lines is never perfectly straight or equally spaced, and the sampling rate is much denser along the flight lines than between them. It is usually desirable to interpolate this flight-line data onto a regular lattice or grid. This procedure is known as *gridding* and permits further algorithms and image processing techniques to be applied to the processed data. Gridding also makes it possible to present the data in the form of an image. Frequently used data presentation techniques are as follows:

Stacked profiles – these are line-based maps in which all flight lines in the data are plotted as XY style graphs. The X axis is along the line of best fit through the flight line, and the Y axis is at right angles to that.

Contour maps – these are maps having contour lines through points of equal anomaly intensities.

Digital images – these are the most common style of presentation today. An example is shown in Figure 2.8. Images are essentially a presentation in which individual pixels in the image are color or gray level coded according to some attribute of the grid points, such as anomaly intensities. The advantage of using images is that they are capable of showing extremely subtle features not apparent in other forms of presentations.

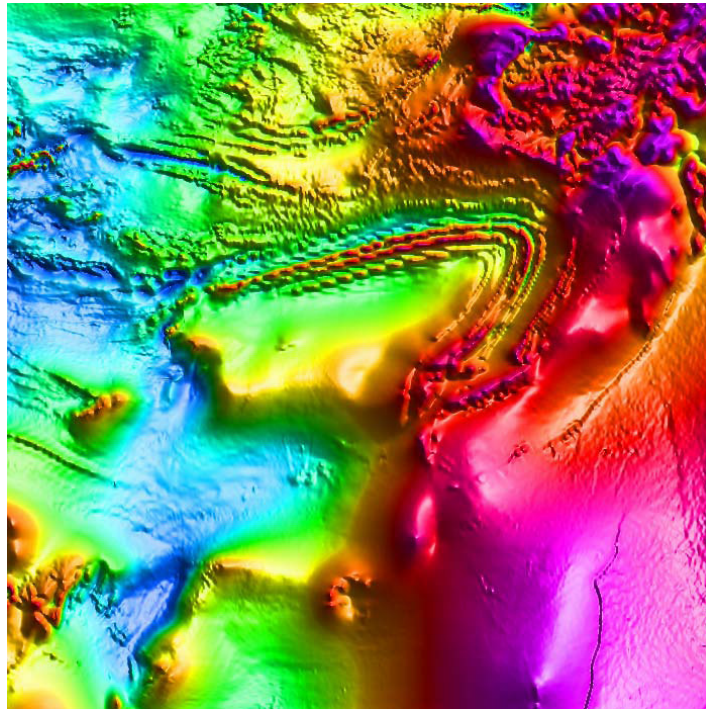


Figure 2.8: Pseudo-colored image of aeromagnetic anomaly survey data (enhanced by artificial illumination) for the Flinders Ranges region of South Australia [13]

Typically, density of the data collected influences the choice of presentation technique. For example, if the flight lines are too far apart, stacked profiles would be the preferred technique as there would not be enough information between the lines to render contour maps or images.

2.6.6 Applications of Aeromagnetic Anomaly Surveys

Aeromagnetic anomaly data, along with various datasets derived through a number of qualitative and quantitative enhancement methods, result in interpretations of the subsurface geology. Most geophysicists provide a qualitative interpretation of the complete dataset with detailed quantitative methods applied to certain anomalies to test the validity of the interpreted source. Qualitative interpretations rely on the spatial patterns that an interpreter can recognize in the data. Faults, dykes, lineaments and folds are usually easily identified [13]. Intrusive bodies are often recognized by virtue of the shape and amplitude of their anomalies. Magnetic lithological units with anomalous susceptibilities can often be directly mapped by recognition of domains with a characteristic magnetic signature. After correlation with additional information, direct lithological inferences can be drawn. Applications that make use of aeromagnetic anomaly surveys are as follows:

1. Exploration of metalliferous minerals and ore deposits
2. Petroleum exploration
3. Assisting geologic mapping
4. Water quality assessment
5. Engineering site surveying
6. Archeological surveys to detect buried metallic objects
7. Environmental audits and site assessments
8. Landfill closures
9. Old pipeline relocation
10. Unexploded ordnance detection

Chapter 3

TEXTURE ANALYSIS

3.1 Definition of Texture

Webster's dictionary defines texture as, "Something composed of closely interwoven elements; *specifically*: a woven cloth: the structure formed by the threads of a fabric" or "The disposition or manner of union of the particles of a body or substance" or "The visual or tactile surface characteristics and appearance of something."

Texture has a *tactile* dimension, i.e., a response felt when a surface is touched and a *visual* dimension, i.e., appearance of distinct patterns or arrangements. In image analysis the focus is on the visual dimension. Visual texture, (henceforth referred to as simply texture) is a descriptor commonly used by human observers in order to interpret pictorial information. It is an innate property of a surface that embodies a structural arrangement or pattern that tends to repeat itself. The arrangement could be periodic, quasi-periodic or random. It is important to note that texture is an organized area phenomenon, a property possessed by a region, which is large enough to display the recurring nature. It implies that texture is sensitive to scale, i.e., the same texture at sufficiently different scales will be perceived as different textures. Researchers have found that the notion of texture doesn't lend itself easily to definition and subsequently to extraction.

Insightful definitions for texture have been proposed in literature and some of them compiled in [20] are listed next.

1. "We may regard texture as what constitutes a macroscopic region. Its structure is simply attributed to the repetitive patterns in which elements or primitives are arranged according to a placement rule."
2. "A region in an image has a constant texture if a set of local statistics or other local properties of the picture function are constant, slowly varying, or approximately periodic."
3. "The image texture we consider is nonfigurative and cellular. An image texture is described by the number and types of its (tonal) primitives and the spatial organization or layout of its (tonal) primitives. A fundamental characteristic of texture: it cannot be analyzed without a frame of reference of tonal primitive being stated or implied. For any smooth gray-tone surface, there exists a scale such that when the surface is examined, it has no texture. Then as resolution increases, it takes on a fine texture and then a coarse texture."
4. "Texture is defined for our purposes as an attribute of a field having no components that appear enumerable. The phase relations between the components are thus not apparent. Nor should the field contain an obvious gradient. The intent of this definition is to direct attention of the observer to the global properties of the display, i.e., its overall "coarseness," "bumpiness," or "finesseness." Physically, non-enumerable (aperiodic) patterns are generated by stochastic as opposed to deterministic processes. Perceptually, however, the set of all patterns without obvious enumerable components will include many deterministic (and even periodic) textures."
5. "Texture is an apparently paradoxical notion. On the one hand, it is commonly used in the early processing of visual information, especially for practical classification

purposes. On the other hand, no one has succeeded in producing a commonly accepted definition of texture. The resolution of this paradox, we feel, will depend on a richer, more developed model for early visual information processing, a central aspect of which will be representational systems at many different levels of abstraction. These levels will most probably include actual intensities at the bottom and will progress through edge and orientation descriptors to surface, and perhaps volumetric descriptors. Given these multi-level structures, it seems clear that they should be included in the definition of, and in the computation of, texture descriptors.”

6. “The notion of texture appears to depend upon three ingredients: (i) some local ‘order’ is repeated over a region which is large in comparison to the order’s size, (ii) the order consists in the nonrandom arrangement of elementary parts, and (iii) the parts are roughly uniform entities having approximately the same dimensions everywhere within the textured region.”

Qualitatively texture can be described as fine, coarse, smooth, granulated, rippled, mottled, irregular, random, lineated, hummocky, etc. These descriptions translate to a set of amplitude (also referred to as gray level or tone or intensity) values and rules for their spatial organization. Haralick [21] observed that amplitude and texture share an inextricable relation. If there is little variation in the amplitude of a given region then the dominant property is the amplitude, whereas if there is wide variation in the amplitude then texture is the dominant property of the region. Hence, descriptors devised to extract texture must recognize the amplitude primitives and their spatial inter-relationship.

3.2 Taxonomy of Texture

Figure 3.1 depicts the taxonomy of texture. Textures, for which a statistical approach is appropriate, i.e., where parameters are determined to estimate a random process or statistical descriptors are used to quantify roughness, are referred to as *disordered* textures. They show neither repetitiveness nor orientation. Textures, for which primitive elements and placement rules that describe the geometrical relation are appropriate, are referred to as *strongly ordered* textures. Textures that are characterized by arbitrary dominant local orientations at each point are referred to as *weakly ordered* or *oriented* textures.

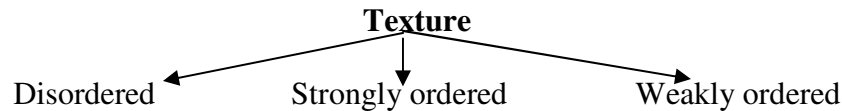


Figure 3.1: Taxonomy of texture [22]

Several quantitative measures have been proposed to characterize and synthesize texture. Two taxonomies of texture measures are described in Figure 3.2 and 3.3.

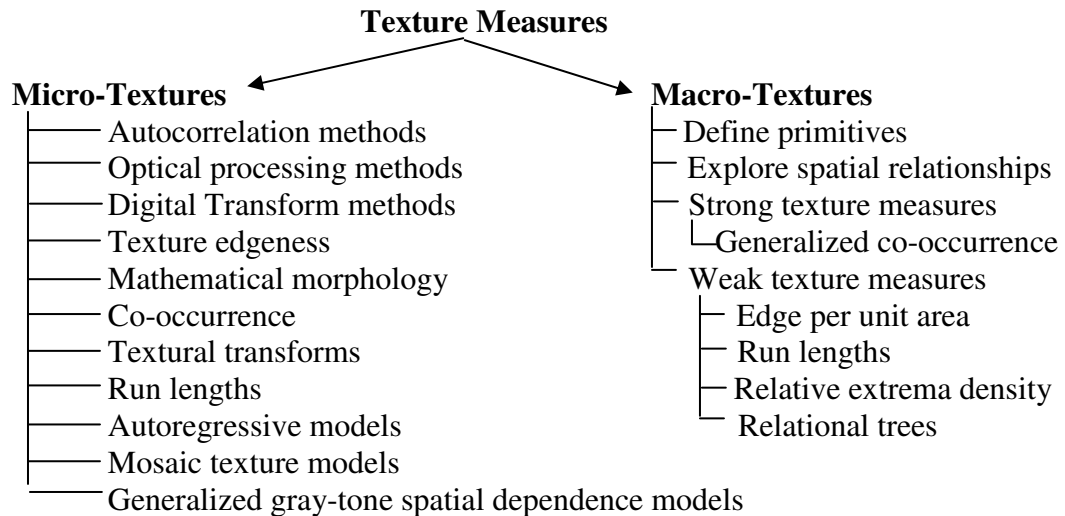


Figure 3.2: Taxonomy of texture descriptors [22]

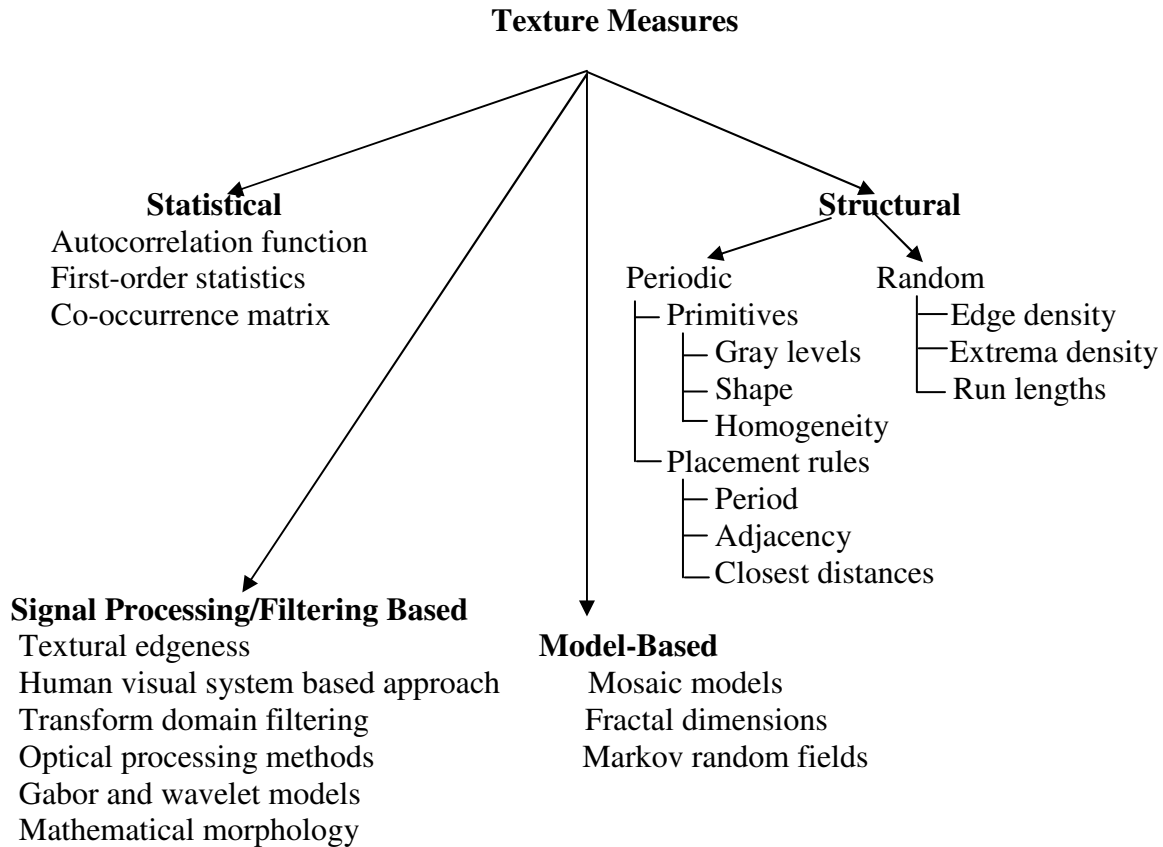


Figure 3.3: Alternate classification of texture measures [14]

It is evident that several sufficiently well defined texture measures are available. The choice of measures to be used depends on the following:

Type of problem – whether it's a texture segmentation, synthesis or classification problem

Image content – whether the image contains strongly ordered, weakly ordered or disordered textures

Prior work in the area – choice of measures previously used and the conclusions drawn about their effectiveness.

Visual inspection of a pseudo-colored aeromagnetic anomaly image reveals that the texture content is highly disordered. There is neither repetitiveness nor orientation. This limits the use of any structural primitives or placement rule based measures. Further,

since the quantitative analysis we undertook was a first of its kind, our choices of statistical measures were driven by consistency with past efforts in the area by Dentith [4] and Gettings [5], as well as a need to limit the scope of the study and keep it simple. We employed five types of texture measures including the following:

1. First-order statistics
2. Second-order statistics
3. Gettings measures
4. Texture spectrum based measures
5. Fractal dimension based measures

A detailed discussion of each of these texture measures follows.

3.3 First-Order Statistics

First-order statistics relate to the likelihood of observing various gray levels (amplitudes or intensities) at a randomly chosen location in the image. These statistics are defined in terms of gray-level histograms. First-order statistics do not take into consideration the relationship with the neighboring pixels. Consider an image (or region of interest) with N pixels, G possible gray levels and n_g pixels of value g . The histogram of the image is defined as follows:

$$h(g) = n_g; \quad g = 0, 1, \dots, G - 1 \quad (3.1)$$

Assuming ergodicity, the normalized histogram measures the probability of a certain pixel occurring in the image. Various first-order statistical measures or moments are derived from the normalized histogram:

$$H(g) = \frac{h(g)}{N}; \quad g = 0, 1, \dots, G-1 \quad (3.2)$$

Mean – value around which the central clustering occurs:

$$\mu = \sum_{g=0}^{G-1} gH(g) \quad (3.3)$$

Variance – variability of the data around the mean:

$$\sigma^2 = \sum_{g=0}^{G-1} (g - \mu)^2 H(g) \quad (3.4)$$

Standard Deviation – square root of the variance.

Coefficient of Variation – ratio of standard deviation to mean:

$$COV = \frac{\sigma}{\mu} \quad (3.5)$$

Skewness – characterizes the degree of asymmetry of a distribution around its mean:

$$Skewness = \sigma^{-3} \sum_{g=0}^{G-1} (g - \mu)^3 H(g) \quad (3.6)$$

Kurtosis – measures the peakness or flatness of a distribution with respect to the normal distribution:

$$Kurtosis = \sigma^{-4} \left\{ \sum_{g=0}^{G-1} (g - \mu)^4 H(g) \right\} - 3 \quad (3.7)$$

3.4 Second-Order Statistics

In his pioneering work on texture discrimination in the 1960s, Julesz [23] attempted to investigate how the human visual system distinguishes between textures having the same brightness, contrast and color. He hypothesized that only if second-order statistics, defined as the likelihood of observing a pair of gray values occurring at the

endpoints of a dipole of random length placed in the image at a random location and orientation, were different would the visual system be able to distinguish between textures. Although Julesz later constructed artificial examples to disprove his conjecture, the work led to interest in measures based on second-order statistics. In their seminal 1973 work, Haralick et al. [17] first proposed the concept of *gray-tone spatial dependency matrices* (also known as *gray-level co-occurrence matrices*) and several novel texture measures that could be extracted from these matrices.

Consider an image with N rows and M columns. Let the domain of row and column elements be $L_x = \{1, 2, 3, \dots, M\}$ and $L_y = \{1, 2, 3, \dots, N\}$, respectively. Let $I(m, n)$ represent the image gray level at pixel (m, n) .

Assuming an 8-connected neighborhood, barring the border pixels, each pixel has two neighbors at each of the angles: $0^\circ, 45^\circ, 90^\circ, 135^\circ$. For each image (or region of interest) four co-occurrence matrices (one for each direction) can be created. For each of the four directions, the co-occurrence matrix is computed by counting the number of occurrences of gray level i , separated by a distance d from gray level j , that occur in the image [17]. Here, “#” indicates the number of elements in the set:

$$P_{0^\circ}(i, j) = \# \left\{ ((k, l), (m, n)) \in (L_x \times L_y) \times (L_x \times L_y) \mid \begin{array}{l} k - m = 0, |l - n| = d, I(k, l) = i, I(m, n) = j \end{array} \right\} \quad (3.8)$$

$$P_{45^\circ}(i, j) = \# \left\{ ((k, l), (m, n)) \in (L_x \times L_y) \times (L_x \times L_y) \mid \begin{array}{l} (k - m = d, l - n = -d) \text{ or } (k - m = -d, l - n = d), I(k, l) = i, I(m, n) = j \end{array} \right\} \quad (3.9)$$

$$P_{90^\circ}(i, j) = \# \left\{ ((k, l), (m, n)) \in (L_x \times L_y) \times (L_x \times L_y) \mid \begin{array}{l} |k - m| = d, l - n = 0, I(k, l) = i, I(m, n) = j \end{array} \right\} \quad (3.10)$$

$$P_{135^\circ}(i, j) = \# \left\{ ((k, l), (m, n)) \in (L_x \times L_y) \times (L_x \times L_y) \mid \begin{array}{l} (k - m = d, l - n = d) \text{ or } (k - m = -d, l - n = -d), I(k, l) = i, I(m, n) = j \end{array} \right\} \quad (3.11)$$

These matrices are symmetric, i.e.

$$P_\theta(i, j) = P_\theta(j, i) \quad (3.12)$$

To illustrate the construction of the co-occurrence matrices, consider a 4×4 image shown in Figure 3.4 (a), where 0, 1, 2, 3 are the gray-level values. Figure 3.4 (b) shows the four co-occurrence matrices for $d = 1$, corresponding to this simple image.

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

Figure 3.4(a): Example image

$P_{0^\circ}(i, j)$

4	2	1	0
2	4	0	0
1	0	6	1
0	0	1	2

 $P_{90^\circ}(i, j)$

6	0	2	0
0	4	2	0
2	2	2	2
0	0	2	0

$P_{45^\circ}(i, j)$

4	1	0	0
1	2	2	0
0	2	4	1
0	0	1	0

 $P_{135^\circ}(i, j)$

2	1	3	0
1	2	1	0
3	1	0	2
0	0	2	0

Figure 3.4(b): For the example image, using $d=1$

Haralick et al. [17] proposed 14 measures to be derived from the co-occurrence matrices. We discuss these measures next.

$p(i, j) = \frac{P(i, j)}{R}$ is the normalized gray-level co-occurrence matrix. R is the number of neighboring resolution cell pairs used in computing a matrix. N_g is the number of possible gray levels in the quantized image. The marginal-probability matrices are

$$p_x(i) = \sum_{j=1}^{N_g} p(i, j)$$

$$p_y(j) = \sum_{i=1}^{N_g} p(i, j)$$

$$p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{\substack{j=1 \\ i+j=k}}^{N_g} p(i, j) \quad k = 2, 3, \dots, 2N_g$$

$$p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{\substack{j=1 \\ |i-j|=k}}^{N_g} p(i, j) \quad k = 0, 1, \dots, N_g - 1$$

The expressions for the 14 texture measures used in this study are as follows:

$$\text{Angular Second Moment:} \quad f_1 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \{p(i, j)\}^2 \quad (3.13)$$

$$\text{Contrast:} \quad f_2 = \sum_{n=0}^{N_g-1} n^2 \sum_{i=1}^{N_g} \sum_{\substack{j=1 \\ |i-j|=n}}^{N_g} \{p(i, j)\}^2 \quad (3.14)$$

$$\text{Correlation:} \quad f_3 = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} i \cdot j \cdot p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (3.15)$$

$$\text{Sum of Squares:} \quad f_4 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu)^2 p(i, j) \quad (3.16)$$

$$\text{Inverse Difference Moment:} \quad f_5 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i, j)}{1 + (i - j)^2} \quad (3.17)$$

$$\text{Sum Averages:} \quad f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i) \quad (3.18)$$

$$\text{Sum Variance:} \quad f_7 = \sum_{i=2}^{2N_g} (i - f_6)^2 p_{x+y}(i) \quad (3.19)$$

$$\text{Sum Entropy:} \quad f_8 = - \sum_{i=2}^{2N_g} p_{x+y}(i) \log\{p_{x+y}(i)\} \quad (3.20)$$

$$\text{Entropy:} \quad f_9 = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log(p(i, j)) \quad (3.21)$$

$$\text{Difference Variance:} \quad f_{10} = \sum_{i=0}^{N_g-1} (i - \mu_{x-y})^2 p_{x-y}(i) \text{ (= variance of } p_{x-y}) \quad (3.22)$$

$$\text{Difference Entropy:} \quad f_{11} = \sum_{i=0}^{N_g-1} p_{x-y}(i) \log\{p_{x-y}(i)\} \quad (3.23)$$

Information Measures of Correlation:

$$f_{12} = \frac{HXY - HXY1}{\max\{HX, HY\}} \quad (3.24)$$

$$f_{13} = (1 - \exp[-2.0(HXY2 - HXY)])^{1/2} \quad (3.25)$$

Here,

$$HXY = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log(p(i, j))$$

$$HXY1 = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log\{p_x(i) p_y(j)\}$$

$$HXY2 = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_x(i) p_y(j) \log\{p_x(i) p_y(j)\}$$

$$\text{Maximum Probability:} \quad f_{14} = \max_{i,j} p(i, j) \quad (3.26)$$

3.5 Gettings Measures

Gettings [5] observed that aeromagnetic anomalies were fractal in nature and highly variable. Based on the example shown in Figure 3.5, he argued that first-order

statistics were not well suited to distinguish between anomalies generated by different sources.

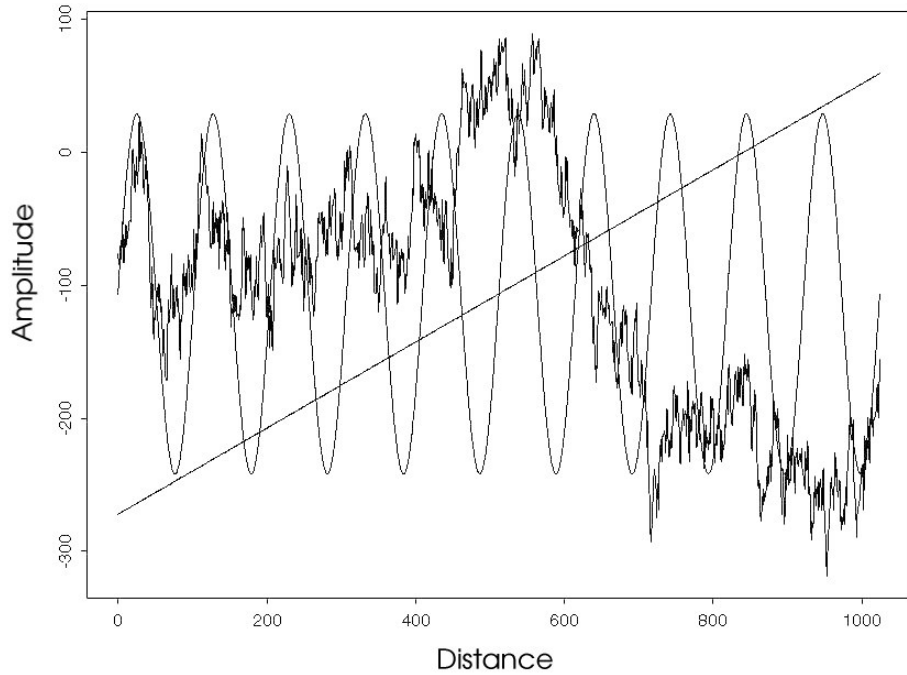


Figure 3.5: Three functions with identical mean and standard deviation [5]

Instead, he proposed two wavelet-based texture measures that could better represent the fractal nature of the data. A discussion of these two measures follows.

Gettings' measures are calculated using a windowing mechanism. The measures are computed using data points encompassed by the window and assigned to the point lying at the center of the window. Using a sliding window, this procedure is repeated along the entire aeromagnetic anomaly flight line. Window size is determined by the size and depth of geological structure being studied. A discussion on the choice of window sizes can be found in Gettings' work [5]. Consider the segment of aeromagnetic anomaly

flight line shown in Figure 3.6. Let x_1, x_2, \dots, x_n , be the positions at which anomaly intensities y_1, y_2, \dots, y_n , are measured. For a window containing w data points, Gettings' measures are defined as follows:

$$\text{Euclidean Length} = \frac{\sum_{i=1}^{w-1} [(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2]^{1/2}}{(x_w - x_1)} \quad (3.27)$$

$$\text{No. of Peaks and Troughs} = \frac{\text{No. of changes in slope in the window}}{(x_w - x_1)} \quad (3.28)$$

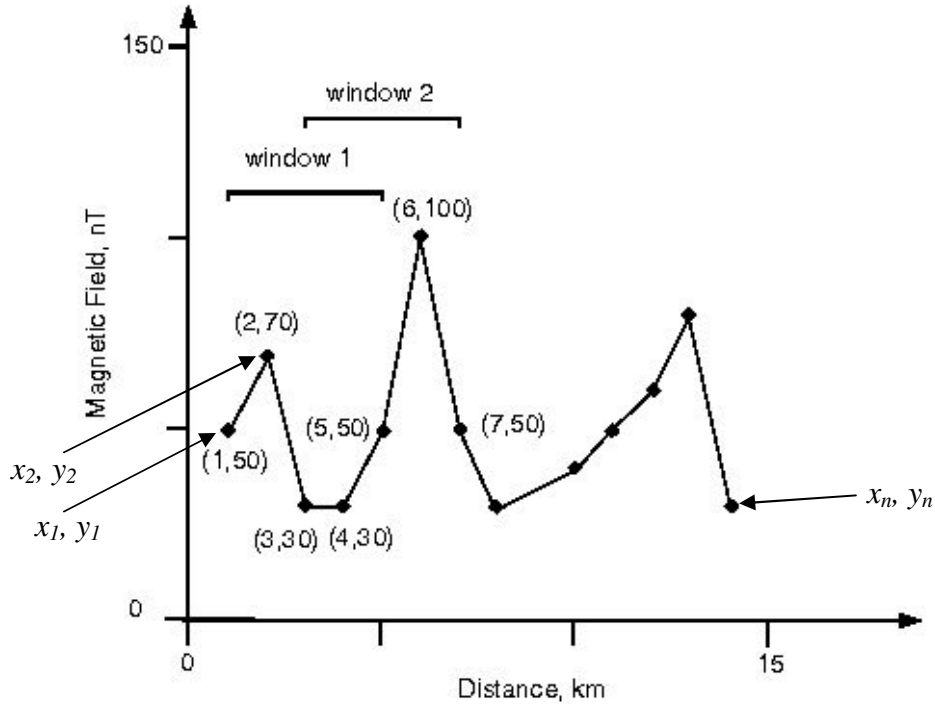


Figure 3.6: Aeromagnetic anomaly flight-line data [5]

Euclidean length is a measure of the amplitude of the data and attempts to capture the intensity of magnetization, whereas number of peaks and troughs captures the undulating nature of the magnetic signal. Unlike GLCM-based measures, these measures

use the raw aeromagnetic anomaly flight-line data itself rather than a re-sampled or quantized version of it. Further, they do not capture textures spread over two dimensions, i.e., in the area between flight lines.

Applying Equations (3.27) and (3.28), we compute Euclidean length and no. of peaks and troughs for windows 1 and 2 (marked in Figure 3.6) as follows:

Window 1: (5 points, 4 km width)

$$\begin{aligned} \text{Euclidean Length} &= \frac{\{[(2-1)^2 + (70-50)^2]^{1/2} + [(3-2)^2 + (30-70)^2]^{1/2} \\ &\quad + [(4-3)^2 + (30-30)^2]^{1/2} + [(5-4)^2 + (50-30)^2]^{1/2}\}}{4 \text{ km}} \\ &= 20.27 \frac{[\text{nT}^2 + \text{km}^2]^{1/2}}{\text{km}} \end{aligned}$$

$$\text{No. of Peaks and Troughs} = \frac{2}{4 \text{ km}} = 0.5 \text{ per km}$$

Window 2: (5 points, 4 km width)

$$\begin{aligned} \text{Euclidean Length} &= \frac{\{[(4-3)^2 + (30-30)^2]^{1/2} + [(5-4)^2 + (50-30)^2]^{1/2} \\ &\quad + [(6-5)^2 + (100-50)^2]^{1/2} + [(7-6)^2 + (50-100)^2]^{1/2}\}}{4 \text{ km}} \\ &= 30.26 \frac{[\text{nT}^2 + \text{km}^2]^{1/2}}{\text{km}} \end{aligned}$$

$$\text{No. of Peaks and Troughs} = \frac{1}{4 \text{ km}} = 0.25 \text{ per km}$$

3.6 Texture Spectrum Based Measures

D.C. He and L. Wang [27] noted that GLCM-based measures depended not only on the spatial relationships of gray levels, but also on the regional intensity variations

within the image. Further, each co-occurrence matrix revealed textural information of the image in a given direction, leaving the user with the problematic choice of which one to use. Keeping these issues in mind, they proposed a new statistical method for texture analysis that focused on texture characterization and discrimination. A *texture unit* was proposed, which is defined as the smallest complete unit that best characterizes the local texture aspect of a given pixel and its neighborhood. Further, the *texture spectrum* of the image was proposed, which describes the distribution of all the texture units within the image. The basic building blocks and measures based on texture spectrum are detailed next.

Given a 3×3 neighborhood in an image, let V_o represent the intensity value of the central pixel and $V_i \{i = 1, 2, \dots, 8\}$ represent the intensity value of the neighboring pixel i . The texture unit is then defined as, $TU = \{E_1, E_2, \dots, E_8\}$, where E_i ($i = 1, 2, \dots, 8$) is determined by the relation,

$$E_i = \begin{cases} 0, & \text{if } V_i < V_o \\ 1, & \text{if } V_i = V_o \\ 2, & \text{if } V_i > V_o \end{cases} \quad (3.29)$$

E_i can also be defined in a binary fashion as follows:

$$E_i = \begin{cases} 0, & \text{if } V_i \leq V_o \\ 1, & \text{if } V_i > V_o \end{cases} \quad (3.30)$$

In the binary format, elements of TU have one of two possible values (0 or 1), so the combination of all eight elements results in $2^8 = 256$ possible texture units.

There is no unique way to label and order the 256 texture units. To achieve that, a *texture unit number* N_{TU} , is used. The texture unit number is defined as follows:

$$N_{TU} = \sum_{i=1}^8 E_i \cdot 2^{i-1} \quad (3.31)$$

Figure 3.7 depicts one possible representation, with the ordering starting at position (a) ($i=1$, $V_i=63$, $E_i=1$) and moving clockwise till position (d) ($i=8$, $V_i=88$, $E_i=1$). The texture unit number, calculated using Equation (3.31), is 198.

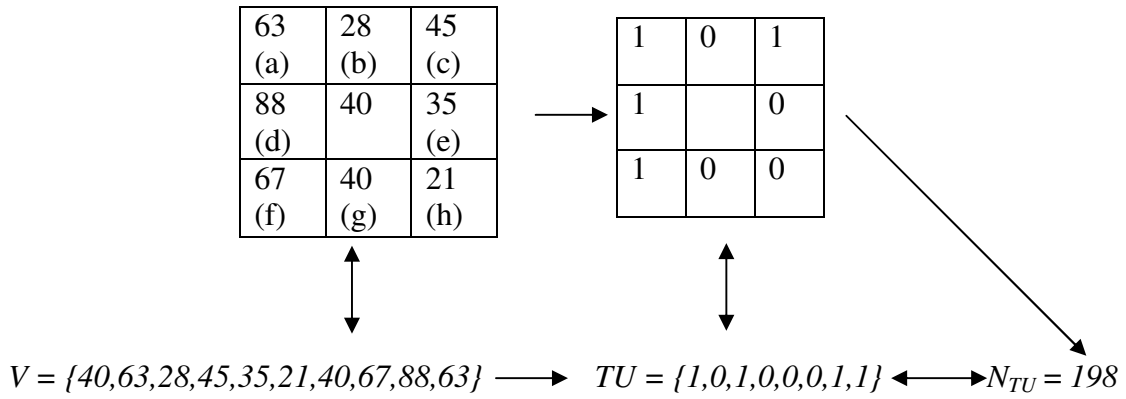


Figure 3.7: Transforming a neighborhood into a texture unit number [27]

The resulting texture unit number is assigned to the center pixel. Once the texture unit number is calculated for every pixel in the image, a frequency distribution of these numbers provides the texture spectrum of the image. The abscissa of the texture spectrum is the texture unit number (N_{TU}) and the ordinate is the frequency of occurrence.

Based on the texture unit octet, texture unit number and the texture spectrum, He and Wang [28] proposed several measures that quantitatively capture texture in an image. Eight texture spectrum based measures used in this study are described next.

Black-White Symmetry (BWS) – measures symmetry between the left and right part of the texture spectrum. BWS is independent of the way in which the texture unit octet is ordered to calculate the texture unit number. Hence, it is an orientation independent measure. Let the texture spectrum be denoted by $S(i)$, where $i=0, 1, 2, \dots, 255$. BWS is defined as follows:

$$BWS = \left[1 - \frac{\sum_{i=0}^{127} |S(i) - S(i+128)|}{\sum_{i=0}^{255} S(i)} \right] \times 100 \quad (3.32)$$

Geometric Symmetry (GS) – measures symmetry in the texture spectrum based on the ordering scheme. It reveals the shape regularity of the image. Let $S_j(i)$ be the texture spectrum under the ordering, where $i=0, 1, 2, \dots, 255$ and $j=1, 2, \dots, 8$ (since there are eight ways the texture unit octet can be ordered). Then GS is defined as

$$GS = \left[1 - \frac{1}{4} \sum_{j=1}^4 \frac{\sum_{i=0}^{255} |S_j(i) - S_{j+4}(i)|}{2 \times \sum_{i=0}^{255} S(i)} \right] \times 100 \quad (3.33)$$

Degree of Direction (DD) – measures degree of linear structure within the image. A high DD value indicates that the texture spectrum is sensitive to the orientation pattern of the image. DD is defined as

$$DD = \left[1 - \frac{1}{6} \sum_{m=1}^3 \sum_{n=m+1}^4 \frac{\sum_{i=0}^{255} |S_m(i) - S_n(i)|}{2 \times \sum_{i=0}^{255} S_m(i)} \right] \times 100 \quad (3.34)$$

BWS, GS and DD are measures based on the geometric appearance of the texture spectrum. They measure the macro-texture of the image. They are insensitive to the

ordering scheme and thus are orientation invariant. The next set of measures attempt to quantify micro-textures.

Micro-Horizontal Structure (MHS) – measures horizontal micro-textures in the image. If $E_a = E_b = E_c$ and $E_f = E_g = E_h$ (Figure 3.11), then MHS is calculated as follows:

$$MHS = \sum_{i=0}^{255} S(i) \times HM(i) \quad (3.35)$$

Here, $S(i)$ is the texture spectrum and $HM(i)$ denotes the horizontal measure of the texture unit numbered i and is calculated as follows:

$$HM(i) = P(a, b, c) \times P(f, g, h) \quad (3.36)$$

$P(a, b, c)$ represents the number of elements having $E_a = E_b = E_c$, while $P(f, g, h)$ is the number of elements with $E_f = E_g = E_h$. A high value of MHS implies the image has strong horizontal micro-structures. Similar measures for micro-textures in the vertical and diagonal directions are defined as follows:

Micro-Vertical Structure (MVS):

$$MVS = \sum_{i=0}^{255} S(i) \times VM(i) \quad (3.37)$$

$$VM(i) = P(a, d, f) \times P(c, e, h) \quad (3.38)$$

Micro-Diagonal Structure (MDS1 and MDS2):

$$MDS1 = \sum_{i=0}^{255} S(i) \times DM1(i) \quad (3.39)$$

$$MDS2 = \sum_{i=0}^{255} S(i) \times DM2(i) \quad (3.40)$$

$$DM1(i) = P(d, a, b) \times P(g, h, e) \quad (3.41)$$

$$DM2(i) = P(b, c, e) \times P(d, f, g) \quad (3.42)$$

Central Symmetry (CS) – is an orientation invariant measure. Let $K(i)$ denote the number of pairs having the same values in elements (E_a, E_h) , (E_b, E_g) , (E_c, E_f) and (E_e, E_d) . CS is defined as follows:

$$CS = \sum_{i=0}^{255} S(i) \times [K(i)]^2 \quad (3.43)$$

3.7 Fractal Dimension Based Measures

Fractals are a class of mathematical functions that attempt to characterize surfaces found extensively in nature. The fractal dimension (FD) of a surface corresponds quite closely to the intuitive notion of *roughness*. A flat 2D plane has a FD=2.0. The rolling countryside could be thought of as having an FD=2.1. An old mountain range (FD=2.3), a young, rugged mountain range (FD=2.5) and finally a stalagmite-covered surface (FD=2.8) are examples of surfaces with increasing fractal dimension [29]. Hence, fractal dimension is a way to characterize surfaces that are not two-dimensional and not quite three-dimensional. Aeromagnetic anomaly data can be visualized as a surface exhibiting roughness. Hence, fractal dimension or some variant of it could prove to be a useful quantitative measure for this data. However, it has been observed that the same fractal dimension can correspond to different fractal sets and virtually identical fractal sets have different fractal dimensions [30]. Such ambiguous properties are highly undesirable in a texture recognition problem. Further, Pilkington [31] and Gettings [4] observed that aeromagnetic anomaly data cannot be characterized by a single fractal dimension. Instead, self-affine fractals or a spectrum of fractal dimensions need to be employed. In spite of these apparent disadvantages, fractal dimension, as a supplementary measure, has

been known to boost performance of other texture descriptors [30]. Procedures used to calculate the two fractal dimension based measures used in this study are discussed next.

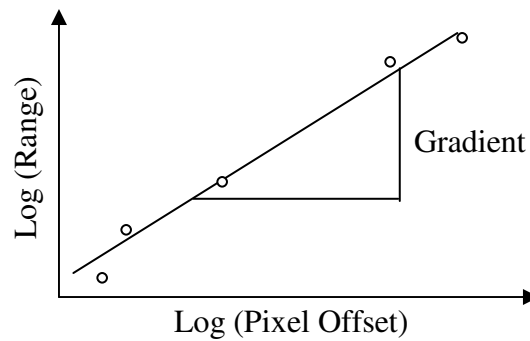
Hurst (fractal) coefficient – is a measure of the fractal dimension of a surface at a given scale. We computed this measure using a method described by Russ [32]. The method is best illustrated by an example. Consider the 5×5 image shown in Figure 3.8.

E	D	C	D	E
D	B	A	B	D
C	A		A	C
D	B	A	B	D
E	D	C	D	E

(a) Example image

	Pixel Location				
	A	B	C	D	E
Pixel Offset	1	$\sqrt{2}$	2	$\sqrt{5}$	$\sqrt{8}$
Range	A_R	B_R	C_R	D_R	E_R

(b) Offset and range calculation



(c) Slope calculation

Figure 3.8: Calculating the Hurst coefficient [32]

Labels ‘A’, ‘B,’ ‘C’, ‘D’ and ‘E’ specify pixels located at distances $1, \sqrt{2}, 2, \sqrt{5}$ and $\sqrt{8}$ respectively, from the center pixel. The first step is to group pixels based on these pixel offsets. Next, the maximum and minimum anomaly intensity values, and in turn the range, is determined for each group. Once these range values are tabulated, the logarithm of range values versus the logarithm of pixel offset is plotted. A regression line is fitted through the points in this plot. The slope of this regression line is known as the Hurst fractal coefficient.

Ratio of Range to Standard Deviation (R/S) – is a normalized measure of variability, which can be used to measure fractal dimension at various scales [5]. It is defined as follows:

$$R / S = \frac{\max - \min}{\sigma} \quad (3.44)$$

Here, σ represents the standard deviation of the values present in a window of data, while *max* and *min* are the maximum and minimum values respectively. In our study, we computed R/S in one-dimensional sliding windows along the flight line, as well as in two-dimensional sliding windows that encompassed multiple flight lines.

Chapter 4

STATISTICAL PATTERN CLASSIFICATION

4.1 Introduction

Humans have the cognitive capability to recognize handwriting, faces, words, shapes, and smells with relative ease. These actions can be grouped under a common act called *pattern recognition*. Pattern recognition can be viewed as a process of taking in sensory data, determining the category (or class) of the pattern observed and taking action based on this knowledge [33]. The advent of digital computers encouraged the development of machines that attempt to automate these tasks. However, the speed and accuracy with which humans are able to perform these acts of recognition belies the complex processes that underlie these abilities. Studies have been conducted to understand the human processes involved in recognizing complex stimulus, and our ability to convert them into meaningful perceptual experiences (such as assigning a name or a sentence to a complex scene). Although a comprehensive model for these processes remains elusive, there is a general agreement on the procedure [34]. First, a pattern is perceived. To represent a meaningful experience, the same pattern or a pattern in the same class should have been perceived previously. Once the past perception is remembered, some equivalence is established between the past and present perceptions, thus allowing for recognition.

Pattern classification is an integral component of pattern recognition. It is the process of assigning a particular pattern to a class. The higher-level reasoning that follows, such as building relationships between the classified patterns, is part of the

complex scene understanding and not included in the classification process. In this study, we restrict ourselves to using automated pattern classification schemes. There is no attempt whatsoever to automate the process of complex scene understanding.

A sub-process of classification is *discrimination*. During discrimination, rules are derived based on patterns whose classes are known. This is followed by the classification stage where these rules are applied to new patterns of unknown classes [35].

A widely accepted sequence of steps for automated pattern classification is depicted in Figure 4.1.

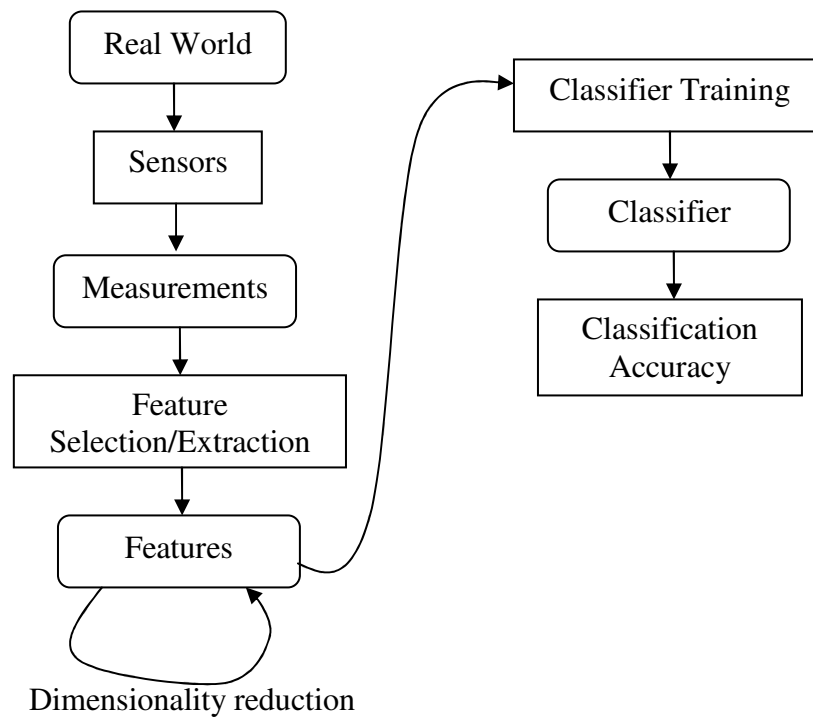


Figure 4.1: Model for automated classification [37]

A sensor, usually some sort of transducer, collects the measurable physical data. The measured data constitutes the universe of samples for the problem at hand. To

simplify subsequent operations without losing any relevant information the measurements may be preprocessed. For example, in our study the raw survey data was uniformly re-sampled to create a more convenient aeromagnetic image.

A feature extractor measures certain properties or features of the data. The goal of the feature extractor is to measure features whose values are very similar for objects of the same category and very different for objects of different categories. Ideally, the features would be distinguishing and invariant to transformations, such as rotation and scale. Several features are extracted to create a multi-dimensional feature vector. Contrary to the intuitive feeling that more features would mean better performance, there is a point of diminishing returns. This phenomenon is referred to as the Bellman's *curse of dimensionality* [34]. Additionally, the features may be redundant and even irrelevant, or the feature vector may be too large to be computationally tractable. A larger feature set also implies more complex classifier logic. These are some of the reasons due to which dimensionality reduction methods are employed. A new feature subset is generated, which serves as input to a classifier. A classifier learning phase follows, involving training based on available samples. Once the training phase is complete, the classifier logic is tested on samples, as yet unseen by the classifier, whose categories are known. This provides the classification accuracy, or the ability of the classifier to correctly recognize new samples. The goal is to maximize classification accuracy, but not at the cost of complicating the classifier logic. This is to ensure that the classifier provides a good generalization.

There are several approaches available to achieve the goal of pattern classification. The important ones include statistical, structural, geometrical, state-space

and neural network approach. In our study, we focused on the statistical approach. The choice of approach was driven by need for simplicity, since this study is the first of its kind and by the need to limit the scope of the study to fit the time available.

In the statistical approach, the pattern-generating source is considered to be a stochastic process [34]. This provides a probabilistic framework to model the relationship between the features and the classes. Multivariate statistical distributions can be used to model the variability in classes. If this approach is employed, the question of whether a pattern belongs to a particular class turns into a special case of the statistical decision theory problem. The statistical framework provides a mathematically rigorous and highly effective platform for the pattern classification problem. However, there are some limitations to this approach. An important one being, the theory of testing statistical hypotheses requires that competing hypotheses be mutually exclusive. This means that a pattern can only be a member of one class. Real-life pattern classification problems usually require more flexibility.

The task of the statistical pattern classifier is to assign a d -dimensional feature vector (a point in d -space), to one out of c possible pattern categories. Two patterns are assigned to the same class based on them being similar and to different classes based on them being dissimilar. This process is aided by features that minimize *within-class* pattern variability and maximize *between-class* pattern variability. Typically, the similarity measures are based on *a priori* probabilities and *likelihood* values of a pattern belonging to a category. These values serve as inputs to some sort of minimum distance classifier, which partitions the feature space so as to minimize the probability of misclassification.

4.2 Classifiers

4.2.1 Bayesian Classifier

Consider the case where there are c possible classes $\omega_1, \omega_2, \dots, \omega_c$ and any given pattern belongs to one of these classes ω_i , with *a priori* probability $P(\omega_i)$, where $P(\omega_i) \geq 0$ and $\sum_{i=1}^c P(\omega_i) = 1$. The feature measurement stage generates a d -component measurement vector or feature vector for each pattern. In the statistical approach, the feature vector \underline{x} is a random vector taking values in a d -dimensional feature space and is governed by the multivariate probability density function $p(\underline{x} | \omega_i)$. Here, $p(\underline{x} | \omega_i)$ is the conditional probability density function of a pattern having values of feature vector \underline{x} and belonging to class ω_i . $p(\underline{x})$ is the probability density function of a pattern having values of feature vector \underline{x} . $P(\omega_i | \underline{x})$ is the conditional probability of a pattern belonging to class ω_i given the fact that the pattern has values of feature vector \underline{x} . This probability is also referred to as the *a posteriori* probability. Typically, in a statistical pattern classification problem, the feature vector \underline{x} is known and the class it belongs to needs to be determined. This can be achieved by computing the *a posteriori* probability $P(\omega_i | \underline{x})$. The Bayes theorem for *a posteriori* probability can be used to compute $P(\omega_i | \underline{x})$ as follows:

$$P(\omega_i | \underline{x}) = \frac{p(\underline{x} | \omega_i)P(\omega_i)}{p(\underline{x})} = \frac{p(\underline{x} | \omega_i)P(\omega_i)}{\sum_{i=1}^c p(\underline{x} | \omega_i)P(\omega_i)} \quad (4.1)$$

If the joint probability density of the pattern generating source, i.e., $p(\underline{x}, \omega_i) = p(\underline{x} | \omega_i)P(\omega_i)$ were known, this would be a trivial problem. However, for most real-life pattern classification problems this information is not available. Instead, an estimate of this probability density function needs to be made based on the available samples.

4.2.2 Bayes Decision Rule for Minimum Error Classification

Consider a classification decision rule $\hat{\omega}(\underline{x}) = \omega_i$, i.e., given a feature vector \underline{x} a decision is taken to assign the feature vector (and in turn the pattern) to class ω_i . To minimize the classification error, the decision rule should be as follows:

$$\hat{\omega}(\underline{x}) = \omega_i \quad \text{if } P(\omega_i | \underline{x}) > P(\omega_j | \underline{x}), \quad j = 1, 2, \dots, c, \quad i \neq j \quad (4.2)$$

Equation (4.2) can also be expressed as

$$\hat{\omega}(\underline{x}) = \omega_i \quad \text{if } P(\omega_i | \underline{x}) = \max_{i=1, \dots, c} P(\omega_i | \underline{x}) \quad (4.3)$$

In certain cases, a reject option could also be provided in which case no decision is made regarding the class of the pattern. In this study, the reject option was not used. Further, the cost (or loss) associated with a correct classification was taken as zero and the cost associated with any wrong classification was taken as one. Under these assumptions, the conditional Bayes risk of choosing class ω_i for pattern \underline{x} is then defined as

$$e(\underline{x}) = 1 - \max_{i=1, \dots, c} P(\omega_i | \underline{x}) \quad (4.4)$$

The average Bayes risk, with the integral extending over the entire feature space, is defined as

$$E = \int e(\underline{x}) p(\underline{x}) d\underline{x} \quad (4.5)$$

Alternatively, the probability of correct classification with the Bayes rule is, $C = 1 - E$. This may be viewed as a measure of the *discriminatory information* conveyed by the features making up pattern \underline{x} .

4.2.3 Discriminant Functions

A classifier can be conceptualized as a network or a machine that computes several *discriminant functions*, $g_i(\underline{x})$, $i = 1, \dots, c$ for a given feature vector \underline{x} , and selects the class based on the largest discriminant. The classifier assigns feature vector \underline{x} to class ω_i if

$$g_i(\underline{x}) > g_j(\underline{x}) \quad \text{for all } j \neq i \quad (4.6)$$

During the training phase, based on the known samples, discriminant functions are used to devise certain decision rules. The effect of these decision rules is that the feature space gets divided into *decision regions*. The regions are separated by *decision boundaries*, which are surfaces in the feature space where ties occur among the discriminant functions. If the classifier's discriminant function is modeled after the *a posteriori* probability described in Equation (4.1), then it is a Bayesian minimum-error classifier. Multiplication by the same positive constant or a shift resulting from the addition of a constant does not change the resulting classification decision. Hence, in case of the Bayesian classifier $g_i(\underline{x})$ could be one of the following:

$$g_i(\underline{x}) = \frac{p(\underline{x} | \omega_i)P(\omega_i)}{\sum_{i=1}^c p(\underline{x} | \omega_i)P(\omega_i)} \quad (4.7)$$

$$g_i(\underline{x}) = p(\underline{x} | \omega_i)P(\omega_i) \quad (4.8)$$

$$g_i(\underline{x}) = \ln p(\underline{x} | \omega_i) + \ln P(\omega_i) \quad (4.9)$$

More generally, $g_i(\underline{x})$ could be any discriminant function that facilitates the decision process. In our study, we used two classifiers whose discriminant functions are based on totally different approaches. The choice of classifiers was driven by the need for simplicity in the analysis and to limit the scope of the study. We discuss these classifiers next.

4.2.4 Gaussian Classifier

The Gaussian classifier is a parametric classifier that attempts to estimate the probability density function $p(\underline{x}, \omega_i) = p(\underline{x} | \omega_i)P(\omega_i)$ from the available samples. Typically, such estimation requires the computation of the parameters that govern the synthesis of the probability density function. The *a priori* probability $P(\omega_i)$ can be estimated by computing the frequency of occurrence of samples in various classes. However, in order to estimate the conditional probability density function $p(\underline{x} | \omega_i)$, an assumption needs to be made as to the nature of the function. One of several standard probability distribution functions could be used to approximate the conditional density. Some of the popular ones include the Gaussian or normal density, exponential density and uniform density. If it is assumed to be Gaussian, then the parametric classifier is cast as a Gaussian classifier. In the case of the Gaussian classifier, $d + d \times (d + 1) / 2$ parameters need to be estimated in order to represent the conditional density as a Gaussian density.

Let the feature vector \underline{x} for a given class ω_i be a continuous valued, randomly corrupted version of a single prototype vector $\underline{\mu}_i$. Then the general multivariate Gaussian conditional density in d dimensions is

$$p(\underline{x} | \omega_i) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp \left[-\frac{1}{2} (\underline{x} - \underline{\mu}_i)' \Sigma_i^{-1} (\underline{x} - \underline{\mu}_i) \right] \quad (4.10)$$

Here, \underline{x} is a d -component column vector, $\underline{\mu}_i$ is the d -component mean vector for class ω_i , Σ_i is the d -by- d covariance matrix for class ω_i and $|\Sigma_i|$ and Σ_i^{-1} are its determinant and inverse respectively. Hence, the conditional probability density function $p(\underline{x} | \omega_i)$, represented concisely by $p(\underline{x} | \omega_i) \sim N(\underline{\mu}_i, \Sigma_i)$, can be estimated by computing d -components of the mean vector $\underline{\mu}_i$ and $d \times (d+1)/2$ unique components of the covariance matrix Σ_i . Samples from a Gaussian density tend to cluster in a single cloud. The center of the cloud is determined by the mean vector and the shape of the cloud by the covariance matrix. In a d -dimensional space the shape of the clusters can be thought of as hyper-ellipsoids.

Using the expression for conditional density in Equation (4.10) and the log-based expression in Equation (4.9), we obtain the discriminant function and decision rule for the Gaussian classifier for minimum error rate classification:

$$g_i(\underline{x}) = -\frac{1}{2} (\underline{x} - \underline{\mu}_i)' \Sigma_i^{-1} (\underline{x} - \underline{\mu}_i) - \frac{d}{2} \ln(2\pi) - \frac{1}{2} \ln |\Sigma_i| + \ln P(\omega_i)$$

$$\hat{\omega}(\underline{x}) = \omega_i \quad \text{if} \quad g_i(\underline{x}) > g_j(\underline{x}) \quad \text{for all } j \neq i \quad (4.11)$$

This representation is also referred to as the quadratic Gaussian classifier due to the quadratic form of the quantity $(\underline{x} - \underline{\mu}_i)' \Sigma_i^{-1} (\underline{x} - \underline{\mu}_i)$. This quantity is sometimes called the squared *Mahalanobis distance* from \underline{x} to $\underline{\mu}_i$.

4.2.5 k -Nearest Neighbor Classifier

Unlike parametric classifiers, a non-parametric classifier does not assume that the probability density function, representing the pattern generating mechanism, has a standard form. A popular non-parametric classifier is the k -nearest neighbor (k -NN) classifier. The idea behind this classifier is that samples that fall close together in feature space are likely to belong to the same class. A simple form of the k -NN classifier is the 1-NN classifier. The decision rule for the 1-NN classifier is as follows:

$$\hat{\omega}(\underline{x}) = \omega_i \quad \text{if} \quad \delta(\underline{x}, \underline{x}_i) = \min_{i=1,2,\dots,n-1} \delta(\underline{x}, \underline{x}_i) \quad (4.12)$$

Here, $\delta(\underline{x}, \underline{x}_i)$ is some metric of the feature space, such as the Euclidian length between the feature vectors \underline{x} and the rest of the feature vectors $\underline{x}_i, i = 1, 2, \dots, n-1$ in the pattern universe. This decision rule partitions the feature space into cells consisting of all points closer to a given point \underline{x} than any other point. The performance of the k -NN classifiers has been found to be surprisingly good as compared to parametric classifiers [33]. However, the storage space and the computational complexity of implementing this classifier are the biggest challenges. Several *editing*, *condensing* or *pruning* techniques have been proposed in order to alleviate these problems. A simple implementation of the editing algorithm [34] used in this study is detailed next. Consider two groups, *condensed* and *total*. To start, let the first feature vector be part of the condensed group and the rest be part of the total group.

1. For each feature vector \underline{x} in total
2. Compute the pre-determined metric to each vector in condensed
3. Find the vector with the minimum metric value and compare the class labels
4. If the class labels are the same do nothing, but if they are different move the vector from total to condensed

5. Repeat this process starting from step 1 till one complete pass is made through the total group and no transfers are made.

This condensing technique results in a significant reduction of samples and greatly reduced the computational complexity without adversely affecting the classification problem at hand.

4.3 Feature Selection and Error Estimation

Although it seems that using more features would result in better classification accuracy, paradoxically there are several motivations for keeping the feature measurements to a minimum. The primary reason is referred to as the “curse of dimensionality” or Bellman’s curse. The problem arises from the fact that as the dimensionality increases the number of samples have to increase commensurately in order to properly compute probability density functions, representing the pattern generating mechanisms, in the multi-dimensional space [33]. In real-life pattern classification problems, such a demand on the number of samples as well as the computation time and complexity is not acceptable. Another reason for feature selection is the removal of redundant or irrelevant information. The presence of such information could definitely have a detrimental effect on the performance of the classifier. Reducing the number of features also makes subsequent analysis easier, faster and relatively more stable.

Rarely are the significant features known *a priori*. Domain knowledge of the pattern classification problem could provide some clues as to the features that will potentially perform better than others. Such a heuristic approach could work. However,

there are several mathematical methods available to achieve this end. Feature selection techniques can be broadly categorized into two main classes, feature selection in the measurement space and feature selection in the transformed space. The latter is also called *feature extraction*.

4.3.1 Feature Selection Criterion and Strategy

Feature selection in the measurement space involves selecting the best subset X of the available d features. Feature selection algorithms are formulated by specifying two ingredients, namely the selection *criterion* and the search *strategy*.

The best subset of features is usually the subset that optimizes some criterion function J . The subset X , must satisfy

$$J(X) = \max_{\theta} J(\theta) \quad (4.13)$$

Here, θ represents the various combinations that arise from the set of d features. The criterion function quantifies the ability of a set of features to achieve maximum class separability. Several criteria are available to make the choice of the best feature subset. The criterion used in this study is called *minimum probability of error*. Theoretically, minimum error probability is computed as stated in Equation (4.5). Algorithmically, it is computed by counting the number of classification errors that arise from the choice of a particular feature.

Ideally, the search strategy for choosing X out of d features would be an optimal search, i.e., the combinatorial formula $\binom{d}{X} = \frac{d!}{(d-X)!X!}$ would apply. However, even for moderate values of d and X , this exhaustive search is not possible. Several sub-optimal search strategies that are more computationally feasible are available. One such

sub-optimal search strategy, called *Sequential Forward Search* (SFS) was used in this study. SFS is a bottom-up search procedure. The SFS algorithm, used in conjunction with the minimum probability of error criterion, is as follows

1. Start with all features in the original set and zero features in the finished subset
2. For every feature in the original set compute the classification error
3. Chose the feature with minimum error, remove it from the original set and place it in the finished subset
4. Take each feature remaining in the original set, combine it with all the feature(s) currently present in the finished subset and compute the classification error
5. Chose the feature that produced the least error in combination with the finished subset, remove it from the original set and place it in the finished subset
6. Repeat steps 4 and 5 until a certain number of features gets selected or until no features remain in the original set

Figure 4.2 depicts the peaking phenomenon used to find the best feature subset.

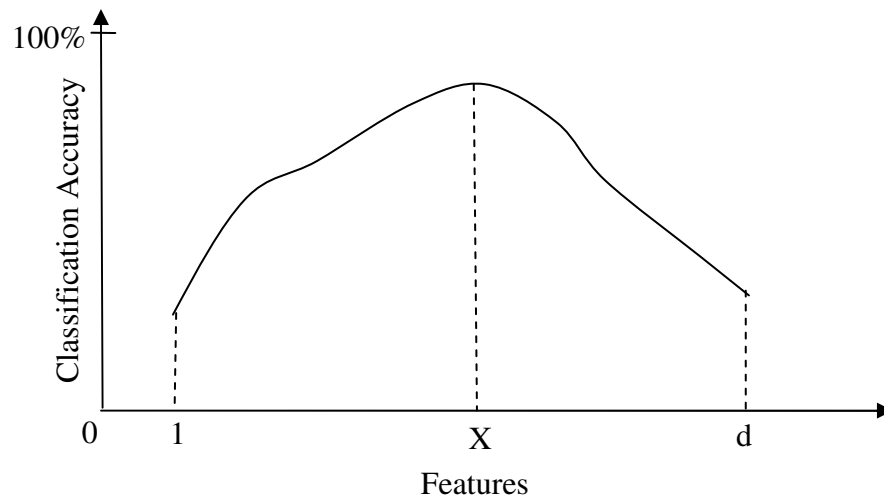


Figure 4.2: Choosing the best feature subset based on peaking accuracy

In our study, we let the SFS algorithm run till there were no features remaining in the original set. Then, using the classification error numbers recorded at every selection step, we tried to find the step where the error recorded was the lowest, i.e., the point where the classification accuracy *peaked*. All features selected till this peak were regarded as the best feature subset.

4.3.2 Feature Extraction

Feature extraction involves a transformation, which uses all the existing dimensions of the feature vectors to yield feature vectors of an equal or lower dimension. This transformation may be a linear or non-linear combination of the original variables. The process of transformation potentially reduces the dimensionality, provides a relevant set of features, reduces the redundancy and ideally produces a lower dimensional representation with enhanced discriminatory information. The *Linear Discriminative Analysis* (LDA) transformation was used in this study. The LDA transformation seeks directions that are efficient for discrimination. Two scatter matrices called *within-class* and *between-class* scatter matrices are computed. Consider a c class problem where each feature vector has d dimensions and assuming $d \geq c$. The within-class matrix is defined as follows:

$$S_w = \sum_{i=1}^c S_i \quad (4.14)$$

$$S_i = \sum (\underline{x} - \underline{\mu}_i)(\underline{x} - \underline{\mu}_i)^t \quad (4.15)$$

S_i is the scatter matrix and $\underline{\mu}_i$ is the mean vector for class ω_i . The between-class scatter matrix is defined as follows:

$$S_B = \sum_{i=1}^c n_i (\underline{\mu} - \underline{\mu}_i)(\underline{\mu} - \underline{\mu}_i)^t \quad (4.16)$$

$\underline{\mu}$ is the mean vector of the entire data and n_i is the number of samples in class ω_i . The principle directions along which the data should be projected, in order to maximize the class discrimination, is obtained by solving for the eigenvectors of the equation

$$(S_B - \lambda_i S_w) \underline{w}_i \quad (4.17)$$

Since the rank of S_B is $c-1$ or less, no more than $c-1$ eigenvalues are non-zero and the weight vectors of \underline{w}_i correspond to these eigenvalues. If W is a d -by- $(c-1)$ matrix with weight vectors \underline{w}_i as its columns, the transformed vectors are calculated as follows:

$$\underline{y} = W^t \underline{x} \quad (4.18)$$

Once the transformation is been complete, the parametric and non-parametric techniques can be applied in the lower-dimensional space.

4.3.3 Error Estimation

In our study, all error estimation was done empirically by counting the number of errors in classification. In the feature selection experiments, error estimation was used as a criterion for choosing a feature. In the classification stage, error estimation was used to compare the performance of various classifiers. Typically, the error was estimated on a validation or test set, completely different from the training set used for classifier learning. However, for a few of the experiments, an independent test set was not available. In those cases, in order to estimate the error, the available samples were split into K parts of equal sizes. $K-1$ parts were used to train the classifier and the error was estimated on the part left out. This procedure was repeated until each part was left out once. This method of error estimation is referred to as *K-fold Rotation* or *K-fold Cross-Validation*.

Chapter 5

DATA AND SOFTWARE

5.1 Data

The different types of data used in this study are described in this section.

5.1.1 Aeromagnetic Anomaly Data

The aeromagnetic anomaly data used in this study was collected and processed by Sial Geoscience Inc. in November 1996, contracted by the U.S. Geological Survey. The survey was conducted over the Patagonia area in Arizona, U.S.A (Region 4191 in Figure 5.1). The data is publicly available [41].

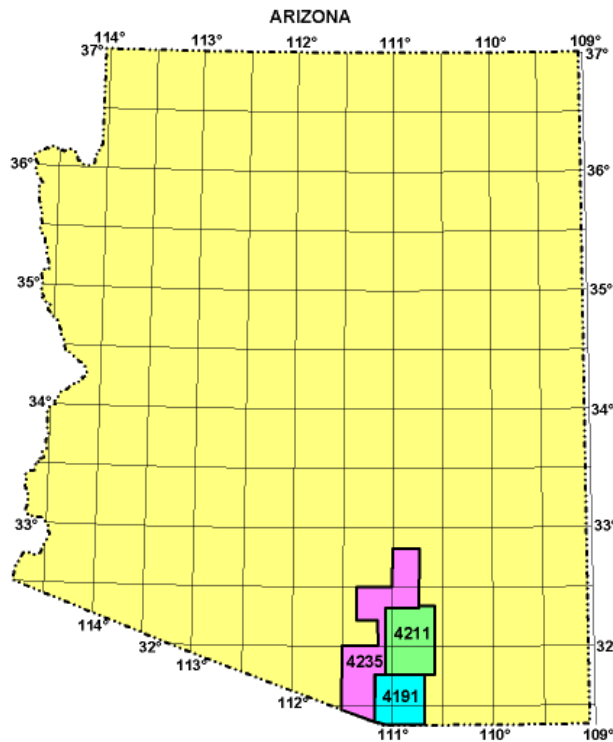


Figure 5.1: Aeromagnetic survey location [41]

The area covered by the survey was approximately 40 km x 40 km (25 miles x 25 miles). The survey was conducted at an average terrain clearance of 225 m. The average distance maintained between the flight lines was 250 m. The average distance between measurements, along the flight line, was 10 m.

The data is published as a flat-file database with approximately one million records. The database is in ASCII format and each record has seventeen fields. These fields are listed in Table 5.1.

1. LINE: Line numbers	10. GPSZ: Barometric altimeter (m)
2. FID: Sample (fiducial) number	11. ALTRM: Radar altimeter (m)
3. FLIGHT: Flight numbers	12. UMAGO: Uncompensated mag (nT)
4. DATE: Year and Julian date	13. MAGO: Compensated mag (nT)
5. TIME: UTC time of day (HHMMSS)	14. BASEB: Diurnal records (nT)
6. GPSX: UTM (m)	15. DRIFT: Diurnal corrections (nT)
7. GPSY: UTM (m)	16. MAGF: Leveled mag (nT)
8. LATITUDE: (degree)	17. MAGIGRF: Residual mag (nT)
9. LONGITUDE:(degree)	

Table 5.1: Fields in the aeromagnetic anomaly database [41]

The Universal Transverse Mercator (UTM) system provides coordinates on a worldwide flat grid [42]. GPSX and GPSY, being UTM coordinates, uniquely identify every point at which the magnetic field is measured. The uncompensated magnetic field (UMAGO) is the total magnetic field at a point due to the rest of the universe. UMAGO mainly consists of the Earth's geomagnetic field, residual magnetic field due to magnetic minerals present directly below the measurement location and the less significant fields

caused by solar, atmospheric and cultural influences. After performing the necessary corrections, the residual magnetic field, MAGIGRF, is computed. This is the *aeromagnetic anomaly* intensity value.

The MAGIGRF values in Patagonia survey database have a large dynamic range of 2500 nT (-1000 nT to +1500 nT) with a measurement accuracy of ± 1 nT.

In order to visualize the data, intensity values are typically rendered as a pseudo-colored image. Figure 5.2 is a pseudo-colored image generated from the aeromagnetic anomaly intensity values measured over the Patagonia area.

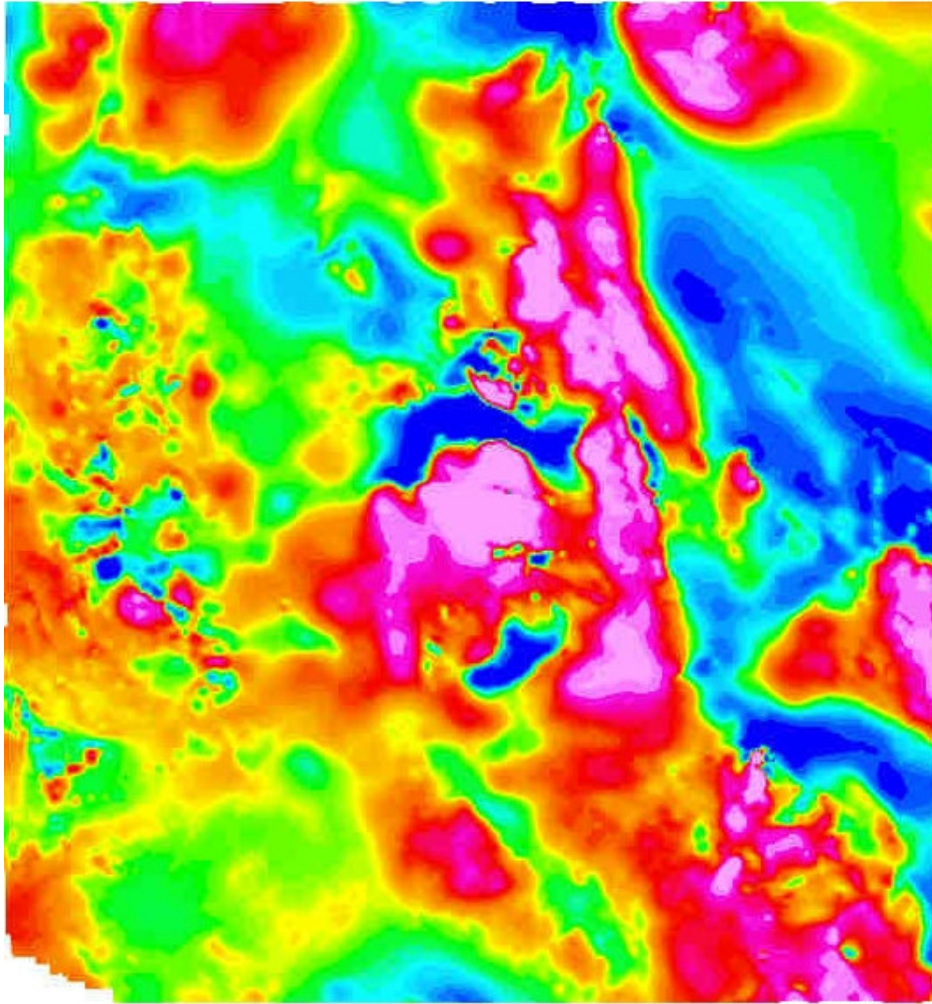


Figure 5.2: Pseudo-colored aeromagnetic anomaly image

A grayscale version of the aeromagnetic anomaly image of the Patagonia region is depicted in Figure 5.3.

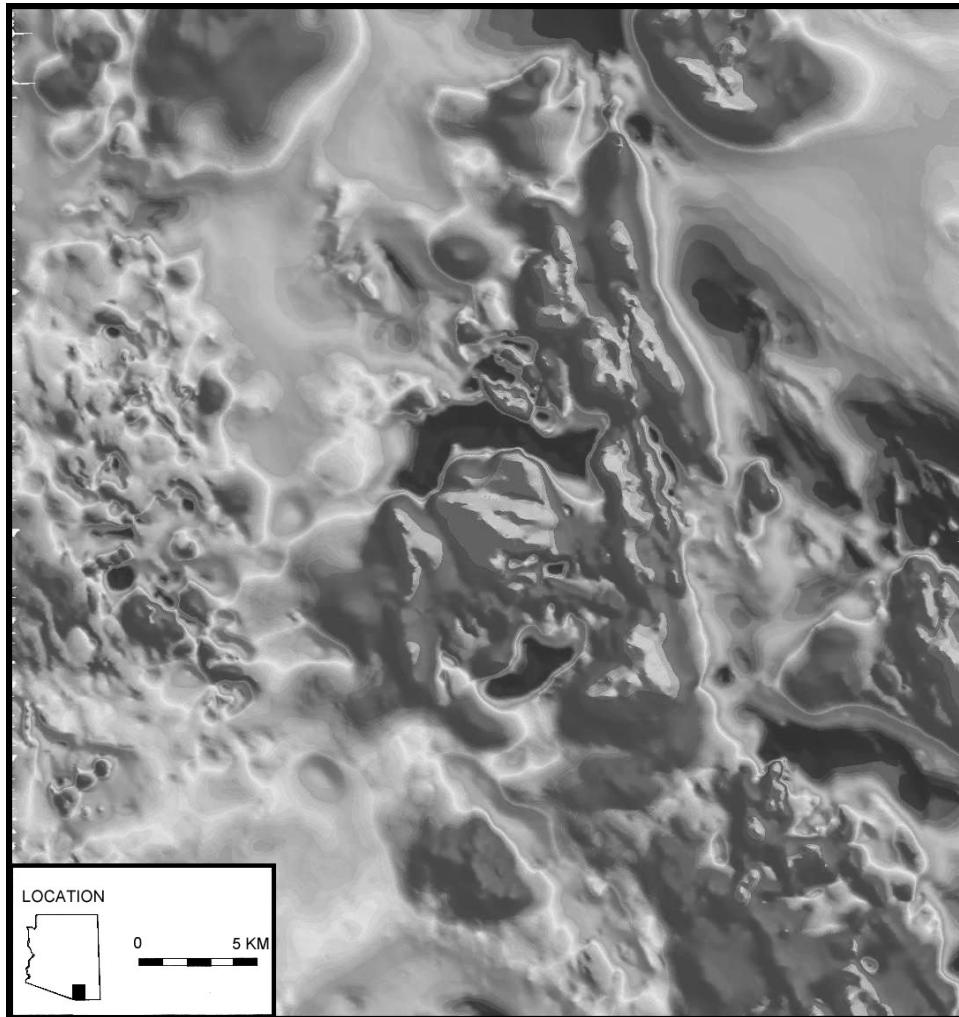


Figure 5.3: Gray-scale version of the aeromagnetic anomaly image [7]

5.1.2 Geologic Data

Geological maps delineate lithological units and geological structures that are present in a geographic region. The southeastern Arizona geological map generated by

Drewes [43], over which the Patagonia aeromagnetic survey was conducted, is shown in Figure 5.4.

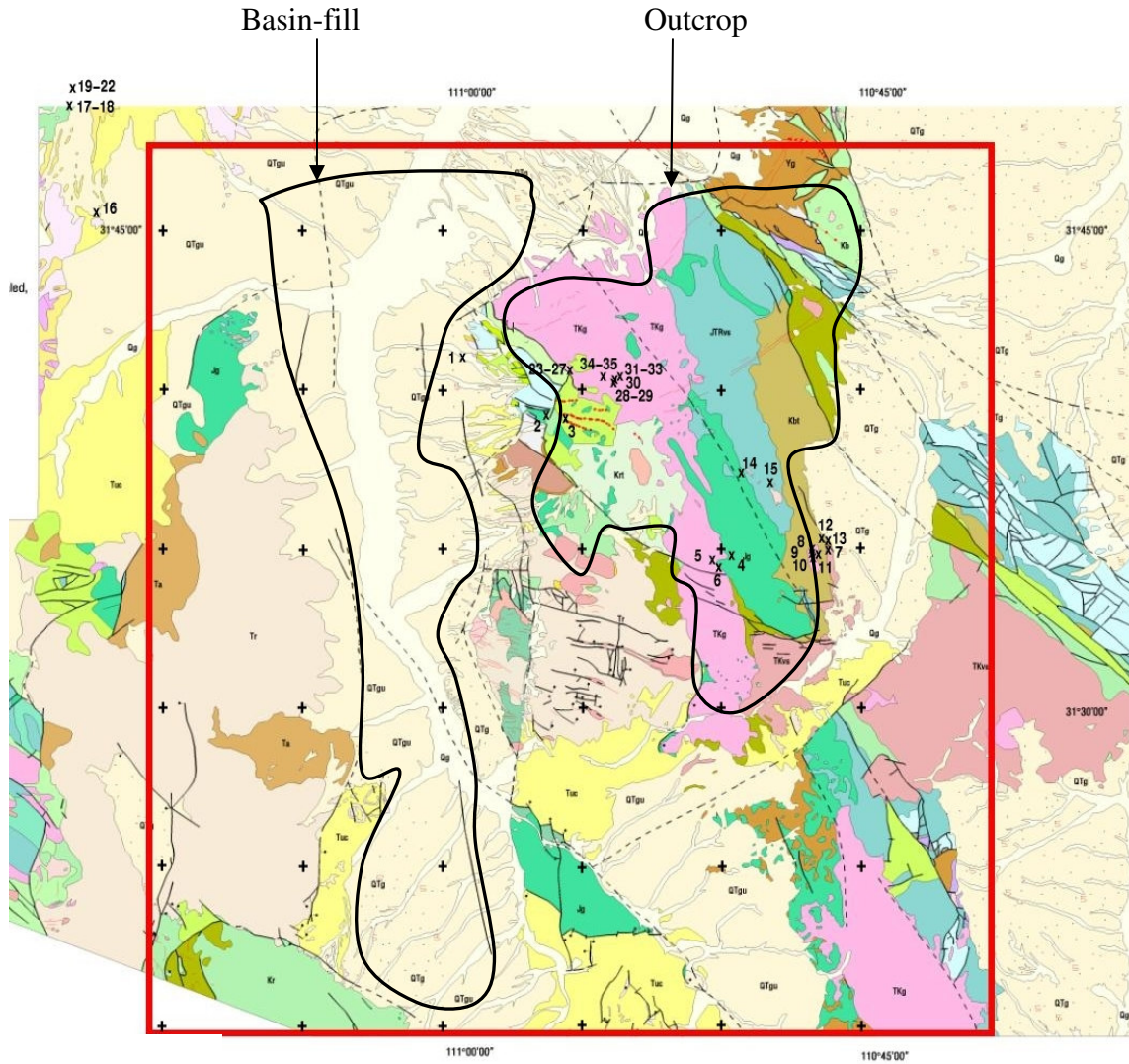


Figure 5.4: Colored geological map of the Patagonia area [4]

(Color indicates lithological units)

The digital version of Drewes' map is publicly available [43]. In our study, we used this digital map at a map scale of 1:125000, to determine the lithological units present in the region.

5.1.3 Composite Aeromagnetic Image

Due logistic difficulties faced during data collection, the aeromagnetic anomaly data provided by the surveyor is not uniformly sampled along the flight line. Further, the distance between the flight lines is not constant either. Applying standard image processing algorithms on non-uniformly sampled data presents a challenge. Hence, a gridding (interpolation) algorithm was employed to create a more convenient image. Details of the gridding technique employed can be found in Appendix A.2. During the gridding process, the sampling interval along the flight lines (*x-axis*) was taken as 10 m and between flight lines (*y-axis*) as 250 m. These choices were based on the sampling intervals targeted by the surveyors while collecting the data.

The gridded values of GPSX, GPSY and MAGIGRF and the lithological unit data from the geologic map were combined to create a composite, henceforth referred to as the *composite aeromagnetic image*. Figure 5.5 depicts the process of creating this image. A detailed discussion of the software packages and procedures used to create the composite image can be found in Appendix A.

The composite aeromagnetic image contains nearly one million pixels. Of these, a little less than half (402,322 pixels) have lithological units that belong to the *outcrop* category and the other half to the *basin-fill* category (marked in Figure 5.4). In outcrop areas, rocks are visible at the surface. Hence the magnetic anomaly can be directly attributed to the rock observed. In basin-fill areas, the rocks are concealed by material such as alluvial fill. The basin-fill material is almost completely non-magnetic. Hence, the magnetic anomaly observed over these areas is due to concealed lithological units. In

these regions the anomaly cannot be unequivocally assigned to any particular lithological unit. So pixels in these regions were not used in the experiments we conducted.

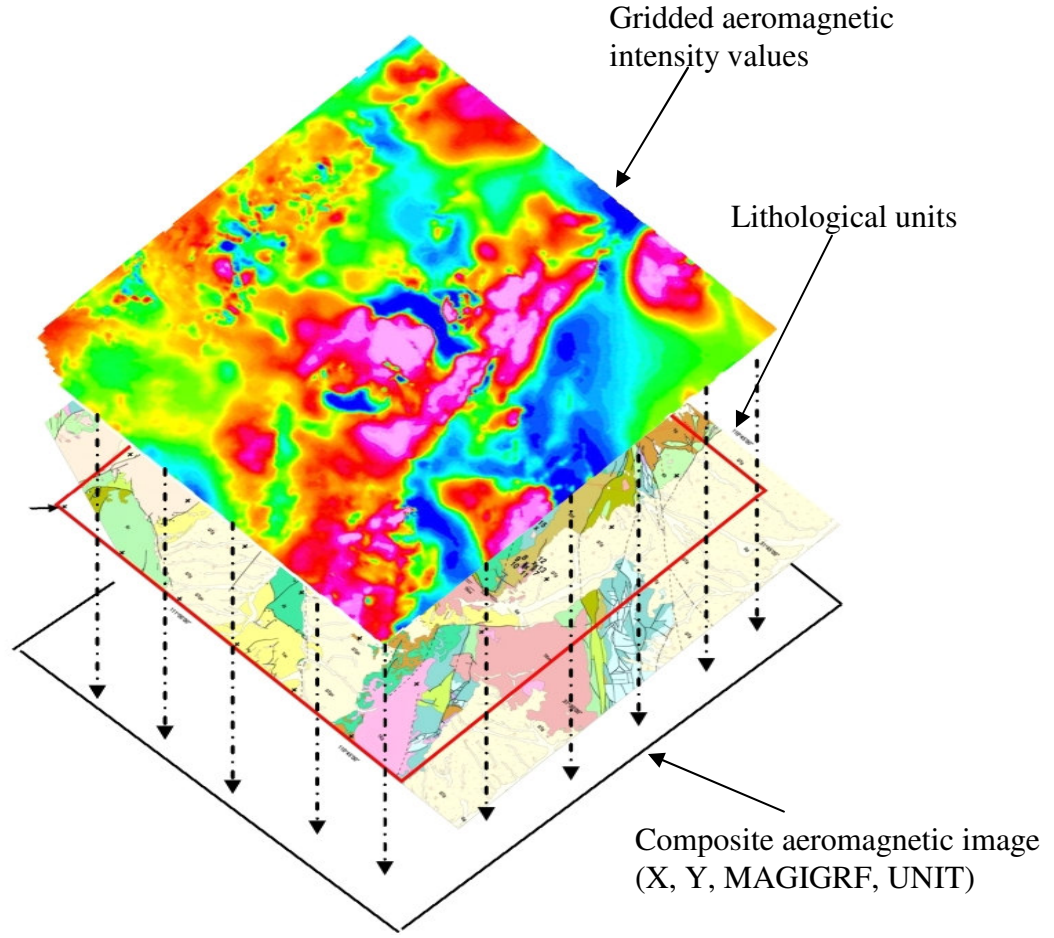


Figure 5.5: Composite aeromagnetic image

In the outcrop region, we found 24 unique lithological units with magnetic mineral content. Of these, 6 units had fewer than 2000 samples. These units had to be eliminated due lack of data, reducing the count of usable lithological units to 18. A comprehensive list of all the lithological units present in the region is included in Appendix B.1.

5.1.4 Brodatz Images

The Brodatz texture image database, containing 32 texture images, was downloaded from the University of Oulu website [47] to test feature measurement code.

5.2 Software

In this section, the different software packages we used and the platforms on which they were operated are described.

5.2.1 Geophysical Package

Oasis montaj is a geophysical data processing software distributed by Geosoft Inc. [44]. We used version 5.1.8 (A5) installed on a Pentium 4, 2.1 GHz, 512 MB RAM system, running Microsoft Windows 2000. The software has several built-in functions that allow the user to grid, filter, re-project, etc. The aeromagnetic survey flat-file database, provided by the surveyor, can be directly loaded into Oasis. We primarily used the software to perform bi-directional gridding and details regarding the steps used can be found in Appendix A.2 and A.3.

5.2.2 Geological Package

ArcGIS is an integrated collection of geographic information systems (GIS) software tools distributed by ESRI Inc. [46]. We used version 8.2 installed on a Pentium 4, 2.1 GHz, 512 MB RAM system, running Microsoft Windows 2000. The software allows users to create, manipulate and view geologic maps. The Patagonia area maps were digitized using this tool. We primarily used the software to intersect the digital map of the Patagonia area with the aeromagnetic survey data to create the composite

aeromagnetic image. The procedure to create the composite image can be found in Appendix A.

5.2.3 Feature Measurement Code

To compute the five types of texture measures, detailed in sections 3.4 to 3.7, we developed an application written in the C programming language. The texture measures were computed on a Pentium 4, 2.26 GHz, 256 MB RAM system, running Red Hat Linux 7.3. The C application was tested to ensure that there were no logical errors in the computation of the texture measures. The test plan for each type of measure is detailed next.

First-order statistics - an artificial image, consisting of 6 rows and 100 columns, was constructed. The six first-order features were computed by hand and using the C application. The results were compared to ensure they were the same.

Second-order statistics – images from the Brodatz texture image database [47] were first split into 32 x 32 sub-images. The sub-images were then randomly split into training and test sets, with each set containing equal number of sub-images from each texture class. For each sub-image, GLCM-based measures were computed using the C application. The measures computed on the training set were provided as input to a Bayesian Gaussian classifier for training. This was followed by an error estimation run on the test set. High accuracy (close to 99%) was achieved in classifying the test set, providing assurance that the GLCM-based measures had been coded correctly.

Gettings measures – using FORTRAN code provided in [5], the Gettings measures were computed for the aeromagnetic data used in this study. When compared with the Gettings measures computed using the C application, they were found to be identical.

Texture spectrum and fractal dimension based measures – the artificial image was used to verify the first-order statistics function, was again employed to compute these measures by hand and the results were compared with the output of the C application. They were found to be identical.

The calculation of 48 measures for 402,322 data points proved to be a computationally intensive task. For the aforementioned hardware and software configuration, Table 5.2 shows the execution time of the C application for certain selected window sizes and quantization levels.

Window Size (m)	Quantization Levels	Execution Time	
		(s)	(hrs, min)
500	16	568	0, 10
500	256	1650	0, 28
1000	256	6175	1, 42
2000	256	17468	4, 51

Table 5.2: Computational complexity for feature measurement code

5.2.4 Pattern Classification Toolbox and Data Visualization

A statistical pattern recognition toolbox, called TOOLDIAG [48], implemented in the C programming language and available as freeware on the web, was employed to perform all the classifier training and error estimation. TOOLDIAG is open source software and includes the implementation of several feature selection, extraction and classification schemes. This software was operated on a Pentium 4, 2.26 GHz, 256 MB RAM system, running Red Hat Linux 7.3. On this system, times taken to execute certain

selected feature selection and classification experiments using TOOLDIAG are recorded in Table 5.3. The recorded experiments were run on a data set of 135,102 data points and feature vector size of 10. The execution times point to the fact that 1-NN classifier experiments took much longer than those involving the Gaussian classifier, while the feature selection step took longer than classification.

TOOLDIAG Experiments	Execution Time (hrs, min)
Classification using Gaussian classifier	0, 2
Classification using 1-NN classifier	4, 35
Sequential forward search, with minimum probability of error criterion using Gaussian classifier	0, 44
Sequential forward search, with minimum probability of error criterion using 1-NN classifier	23, 4

Table 5.3: Computational complexity for TOOLDIAG experiments

In order to visualize the data, exploratory data analysis tools were used. *Scatter plots* use two variables and depending on the plot provide information regarding the relationship between the two variables. They also display the natural grouping or clustering of the data. This second property is a visual indicator on whether class separability can be achieved easily. If there are several classes or there if there is too much confusion regarding the spread of the classes, a *convex hull* could be used as a rough estimate of the extent. The convex hull behaves like a rubber band around the outside points of the class and makes it easier to understand the overlap between classes. Another data display tool suggested by Fukunaga [39] is the d^2 -display. It is a tool best

suited for a two class problem. The Mahalanobis distances of every feature vector to the two classes form the axes. The display provides a visual method of determining the separation between the classes.

The visualization was performed using built-in MATLAB functions, such as *gscatter* and *convhull* and scripts written in the MATLAB programming language.

Chapter 6

EXPERIMENTS AND RESULTS

Reiterating the scope of this study, a composite aeromagnetic image was created for the Patagonia region of southeastern Arizona. Feature measurement algorithms were applied to generate a 49 dimensional feature vector (Appendix C) for each of the 402,322 data points in the outcrop regions of the survey area. 18 lithological units were identified as having sufficient samples (Appendix B), making it an 18-class, 49-dimension pattern classification problem. The broad objective of this study was to determine the quantitative capability of texture measures, extracted from the aeromagnetic anomaly data, in predicting the lithology of a region. Additionally, a quantitative comparison was sought between the various texture measures.

6.1 Experiments Using Survey and Gridded Data

Table 6.1 provides the list of 18 lithological units and the *a priori* probabilities of their occurrence. The units are not equiprobable. Unit 1050 has the highest *a priori* probability of 28.77%. Hence, the classification accuracy numbers were compared to this statistic. We analyzed the raw aeromagnetic intensity data obtained from the survey and its quantitative ability in predicting the lithology of the region. Table 6.2 shows the results for this one-feature classification experiment. The result was no better than the *a priori* probability itself, hence, justifying the need for better discriminating features. To create the composite aeromagnetic image, the raw data is gridded. An experiment was conducted to study the effect of gridding on the classification accuracy.

Lithological Unit	Samples	<i>A Priori</i> Probability (%)
1048	11811	2.94
1050	115738	28.77
1052	4519	1.12
1056	33345	8.29
1058	24534	6.10
1061	31414	7.81
1062	17169	4.27
1063	5173	1.29
1065	27969	6.95
1067	7063	1.76
1080	16778	4.17
1082	42627	10.60
1085	9475	2.36
1086	6673	1.66
1089	28855	7.17
1105	7173	1.78
1110	8547	2.12
1113	3369	0.84

Table 6.1: Summary of lithological units

Classifier	Accuracy (%)	Improvement Over <i>A Priori</i> Probability (%)
Quadratic Gaussian	28.23	-0.54
1-NN	14.64	-14.13

Table 6.2: Classification accuracy for survey intensity values

A comparison of results in Tables 6.2 and 6.3 suggest that the gridding process did not adversely affect the accuracy numbers, giving us confidence in the interpolated data.

Classifier	Accuracy (%)	Improvement Over <i>A Priori</i> Probability (%)
Quadratic Gaussian	27.94	-0.83
1-Nearest Neighbor	13.42	-15.38

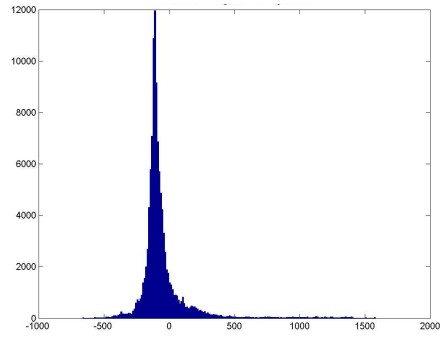
Table 6.3: Classification accuracy for gridded intensity values

6.2 Classification Using Complete Feature Vector

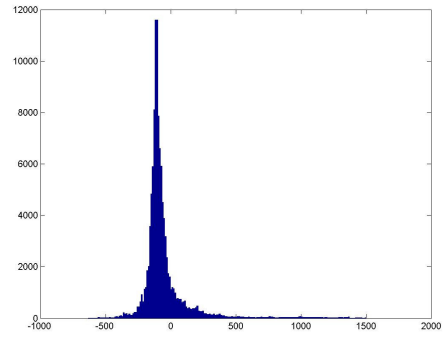
To extract the 48 features for each pixel in the composite aeromagnetic image, a sliding window is used. The rule of thumb is that window size should be greater than the size of the smallest geological structure of interest. To start with, we used a 500 m \times 500 m window. This also ensured that at least 3 flight lines were taken into account during feature measurement. Although, geographically speaking the window is a square, pixel-wise the window is rectangular due to the unequal sampling intervals along the x (10 m) and y (250 m) axes. Computation of the GLCM-based features was preceded by a uniform quantization step. As a starting point, 256 quantization levels were employed. In all, including the raw aeromagnetic anomaly intensity value, a 49 dimensional feature vector was computed for every point in the composite aeromagnetic image.

6.2.1 Class Conditional Feature Histograms

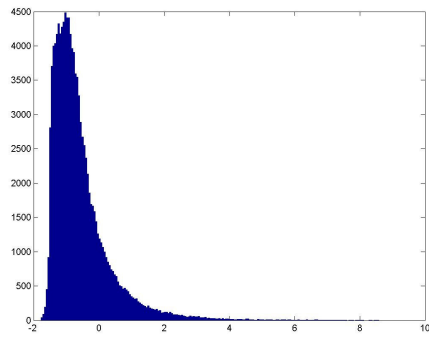
Experiments with the entire set of 49 features were preceded by a visualization exercise. For 3 classes with the highest *a priori* probabilities, namely units 1050, 1056 and 1082, histograms were plotted for 6 features. This exercise provided insight into the shape of the class conditional distribution of the features. Figures 6.1, 6.2 and 6.3 suggest that several features have a unimodal, if not normal, distribution. This provides a qualitative justification for using a Gaussian classifier.



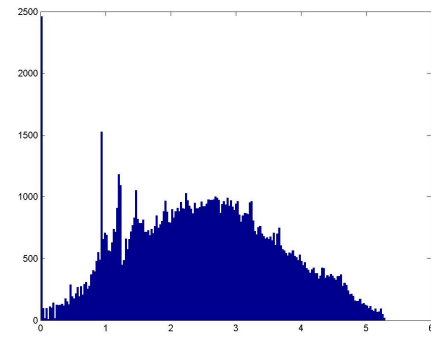
(a) Intensity value



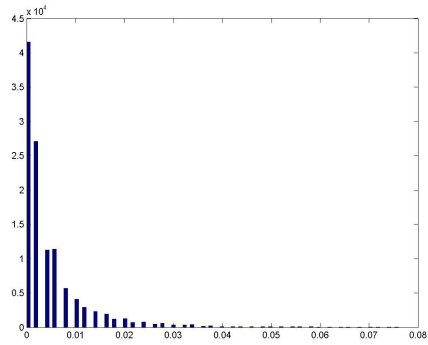
(b) First-order statistic: Mean



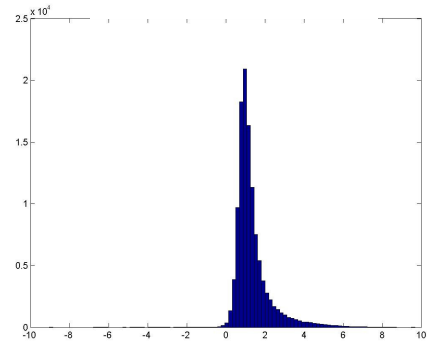
(c) First-order statistic: Kurtosis



(d) Second-order statistic: Entropy

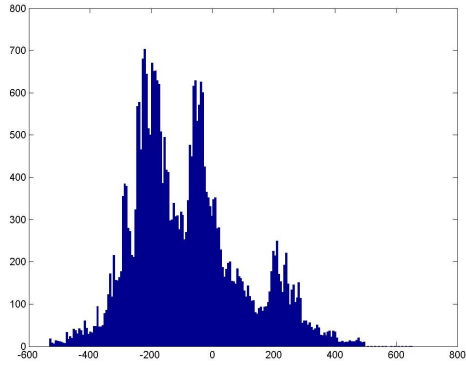


(e) Gettings-measure: Peaks and troughs

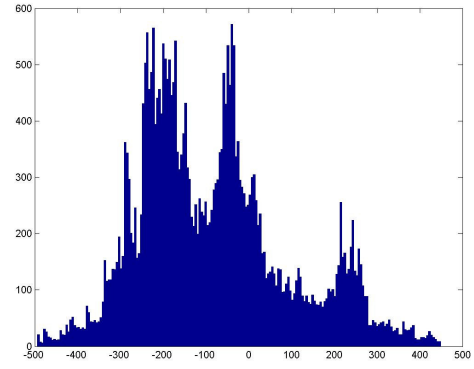


(f) Fractal based: Hurst Coefficient

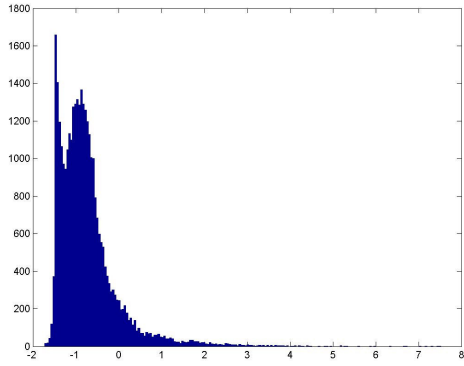
Figure 6.1: Class conditional feature histograms for class 1050



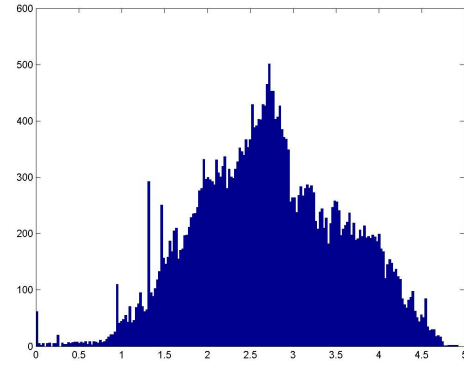
(a) Intensity value



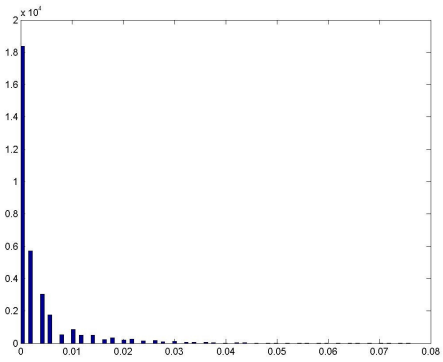
(b) First-order statistic: Mean



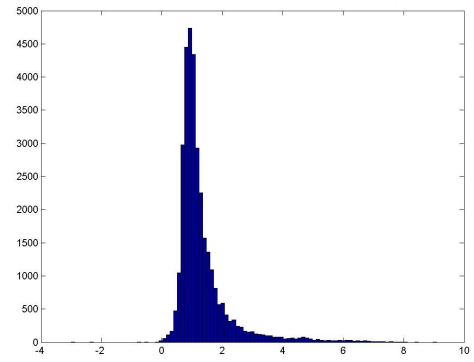
(c) First-order statistic: Kurtosis



(d) Second-order statistic: Entropy

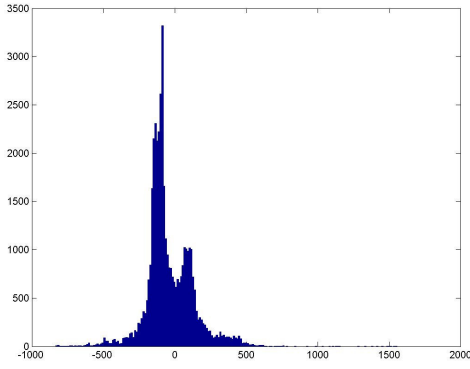


(e) Gettings-measure: Peaks and troughs

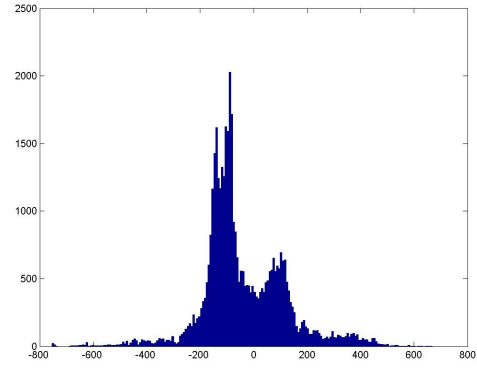


(f) Fractal based: Hurst Coefficient

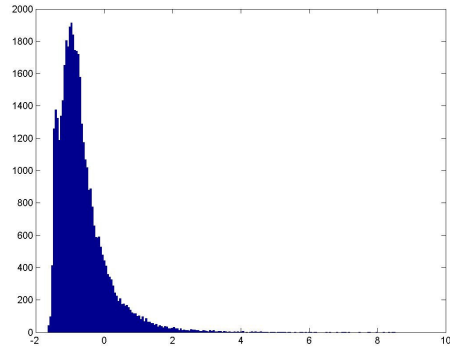
Figure 6.2: Class conditional feature histograms for class 1056



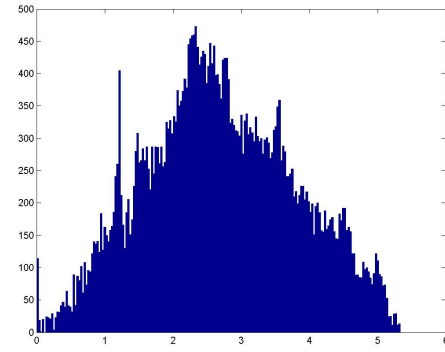
(a) Intensity value



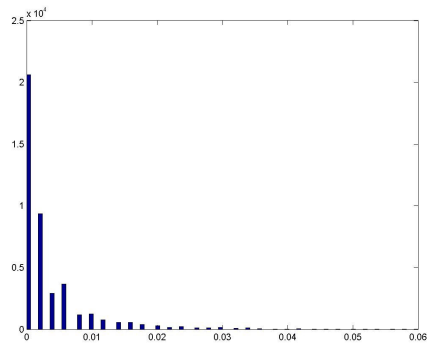
(b) First-order statistic: Mean



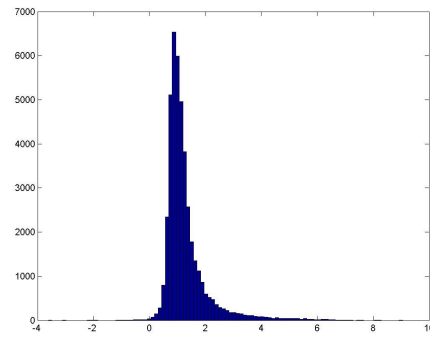
(c) First-order statistic: Kurtosis



(d) Second-order statistic: Entropy



(e) Gettings-measure: Peaks and troughs



(f) Fractal based: Hurst Coefficient

Figure 6.3: Class conditional feature histograms for class 1082

6.2.2 Accuracy Results for Complete Feature Vector

Table 6.4 shows results for the classification experiments with all 49 features. The accuracy numbers were obtained by running three iterations of two-fold cross-validation. The 1-NN classifier outperformed the Gaussian classifier by a large margin. As compared to the accuracy numbers for a single feature (Table 6.3), the accuracy numbers obtained by the Gaussian classifier in this experiment were lower. This implied that there were redundant features that were having a detrimental effect on the classification accuracy. Hence, a feature selection stage was required as a preceding stage to the classifier.

Classifier	Accuracy (%)	Improvement Over A <i>Priori</i> Probability (%)
Quad. Gaussian	11.66	-17.11
1-NN	98.33	69.56

Table 6.4: Classification accuracy for 49 features

6.3 Using Feature Selection Strategies

Two feature selection strategies were used in this study. Feature selection in the measurement space was achieved using Sequential Forward Search (SFS) with minimum probability of error as the criterion function (Section 4.3.1). Feature selection in the transformed space, or feature extraction, was achieved using the Linear Discriminant Analysis (LDA) transformation. Since the number of classes (18) was less than the number of features (49), the LDA transformation generated a subspace of features (less than or equal to 17). Hence, applying the same classifiers after the LDA transformation was not a redundant exercise. Error estimation was conducted using three iterations of two-fold cross-validation. The results, listed in Table 6.5, indicate that the classification

accuracy for the Gaussian classifier improved due to reduction of dimensionality. However, the 1-NN classifier continued to perform much better than the Gaussian classifier. The LDA transformation did boost the performance of the Gaussian classifier but not significantly. The transformation had an adverse effect on the 1-NN classifier performance.

Feature Selection/ Classifier	Accuracy (%)	Improvement Over A <i>Priori</i> Probability (%)
SFS/Quad. Gaussian	33.10	4.33
SFS/1-NN	97.52	68.75
LDA/Quad. Gaussian	35.20	6.43
LDA/1-NN	68.11	39.34

Table 6.5: Classification accuracy after feature selection

6.3.1 Issues with Experimental Setup

Intuitively, the large difference between the accuracy numbers of the 1-NN and the Gaussian classifiers seemed unnatural. It was observed that this difference could be attributed to the methodology used in conducting the experiment.

During the feature measurement process, a sliding window is run across the data to compute the texture measures. Hence, points in the dataset that are spatially close will end up having very similar feature vectors. This is because, in the case of neighboring pixels, the majority of the pixels taken into consideration by the sliding window while computing the texture measures are the same. Further, in the two-fold cross-validation scheme for error estimation, the data is randomly split into two parts. At a time one part is used as the training set and the error estimation is conducted on the other set. The average of the two runs is reported. In randomly splitting the dataset, it was possible that

points in the dataset that were spatially close to each other were being split. Some points were becoming part of the training set while others were part of the test set. This was potentially introducing a bias in the error estimation process. In order to avoid this bias, spatially uncorrelated training and test sets had to be created.

6.3.2 Experiments Using Non-Overlapping Window

A simple way of avoiding aforementioned bias is by using a non-overlapping window instead of a sliding window. This would ensure that the data points did not share the pixels that were being used to compute the features. The downside to this method was that the pattern universe shrunk dramatically. Table 6.6 and Table 6.7 provide the details for this experiment. The reduction in the accuracy numbers for the 1-NN classifier certainly indicated that some sort of bias was present in the previous experiments. However, the drastic reduction in the number of samples could also have led to this drop in classification accuracy.

6.4 Experiments with Spatially Uncorrelated Sets

6.4.1 Experiments with Two Spatially Uncorrelated Sets

In order to remove the spatial bias in the training and test sets while maintaining sufficient number of points, the dataset was manually split into two spatially disjoint sets while maintaining equal *a priori* probabilities for each class across the two sets. Table 6.8 shows the split dataset. Feature selection and classifier training was conducted on set 1, while the classification accuracy results were computed on the spatially independent set 2.

Lithological Unit	Samples	Probability (%)
1048	79	3.00
1050	773	29.37
1052	30	1.14
1056	207	7.68
1058	159	6.04
1061	199	7.56
1062	114	4.33
1063	34	1.29
1065	190	7.22
1067	45	1.71
1080	115	4.37
1082	269	10.22
1085	60	2.28
1086	42	1.60
1089	190	7.22
1105	47	1.79
1110	58	2.20
1113	21	0.80

Table 6.6: Summary of lithological units for the non-overlapping window experiment

Feature Selection/ Classifier	Accuracy (%)	Improvement Over A <i>Priori</i> Probability (%)
SFS/Quad. Gaussian	30.70	1.33
SFS/1-NN	29.43	0.06
LDA/Quad. Gaussian	32.94	3.57
LDA/1-NN	27.20	-2.17

Table 6.7: Classification accuracy after feature selection for non-overlapping window

Lithological Unit	Set 1		Set 2	
	No.	(%)	No.	(%)
1048	5901	2.94	5910	2.94
1050	58515	29.11	57223	28.44
1052	2107	1.05	2412	1.20
1056	16592	8.25	16753	8.33
1058	11816	5.88	12718	6.32
1061	15956	7.94	15458	7.68
1062	8260	4.11	8909	4.43
1063	2598	1.29	2575	1.28
1065	14256	7.09	13713	6.81
1067	3803	1.89	3260	1.62
1080	8165	4.06	8613	4.28
1082	20643	10.27	21984	10.92
1085	5124	2.55	4351	2.16
1086	3416	1.70	3257	1.62
1089	14252	7.09	14603	7.26
1105	3628	1.80	3545	1.76
1110	4074	2.03	4473	2.22
1113	1892	0.94	1477	0.73

Table 6.8: Summary of lithological units for 2 ways split

A possible reason for the low accuracy numbers in Table 6.9 is that the classifier may have been over fit to set 1 resulting in a loss of generalization.

Feature Selection/ Classifier	Accuracy (%)	Improvement Over A <i>Priori</i> Probability (%)
SFS/Quad. Gaussian	28.40	-0.04
SFS/1-NN	27.52	-0.92
LDA/Quad. Gaussian	29.89	1.45
LDA/1-NN	21.52	-6.92

Table 6.9: Classification accuracy for set 2

6.4.2 Experiments with Three Spatially Uncorrelated Sets

In order to eliminate the problems of spatial correlation bias, small sample size and over fitting, a scheme was devised where in the data set was manually split into 3 spatially uncorrelated sets (Table 6.10) with equal *a priori* probability for classes across the sets. Set 1 was used to train the classifier. Set 2 was used for the purpose of feature selection so that the features selected were not over fit to the training data. Set 3 was finally used as an independent test set to get the accuracy estimate.

Lithological Unit	Set 1		Set 2		Set 3	
	No.	(%)	No.	(%)	No.	(%)
1048	4035	3.01	3999	2.96	3777	2.83
1050	38139	28.49	39218	29.05	38381	28.78
1052	1487	1.11	1472	1.09	1560	1.17
1056	10993	8.21	11112	8.23	11240	8.43
1058	8228	6.15	8055	5.97	8251	6.19
1061	10363	7.74	10466	7.75	10585	7.94
1062	5365	4.01	5845	4.33	5959	4.47
1063	1755	1.31	1694	1.25	1724	1.29
1065	9566	7.15	9245	6.85	9158	6.87
1067	2298	1.72	2614	1.94	2151	1.61
1080	5578	4.17	5649	4.18	5551	4.16
1082	14314	10.69	14481	10.72	13832	10.37
1085	3225	2.41	2999	2.22	3251	2.44
1086	2472	1.85	2146	1.59	2055	1.54
1089	9474	7.08	9857	7.30	9524	7.14
1105	2335	1.74	2381	1.76	2457	1.84
1110	2924	2.18	2744	2.03	2879	2.16
1113	1306	0.98	1046	0.77	1017	0.76

Table 6.10: Summary of lithological units for 3 ways split

This type of data split ensured there were no biases. Table 6.11 reflects the unbiased capability of the texture measures in predicting lithological units. The elimination of potential problems in the experimentation methodology did not result in any signification gain in classification accuracy. In fact, the only conclusion from the results was that regardless of the number or type of texture features used, the machine classification algorithms could do no better than *a priori* probability. This difficult predicament led to the conclusion that data visualization was required in order to investigate the potential reasons due to which the machine classification algorithms were performing poorly.

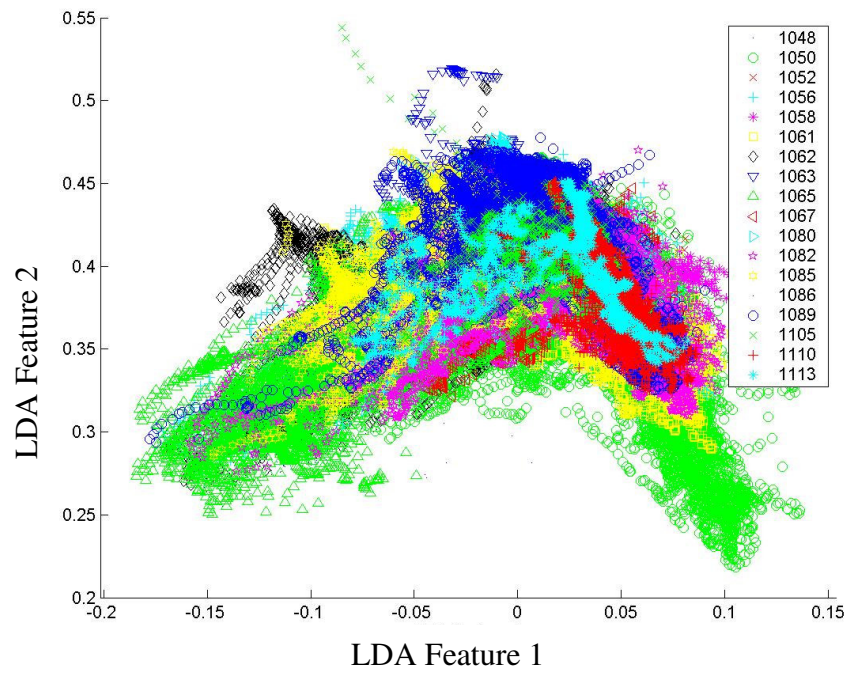
Feature Selection/ Classifier	Accuracy (%)	Improvement Over A <i>Priori</i> Probability (%)
SFS/Quad. Gaussian	28.97	0.19
SFS/1-NN	26.95	-1.83
LDA/Quad. Gaussian	29.23	0.45
LDA/1-NN	22.02	-6.76

Table 6.11: Classification accuracy for set 3

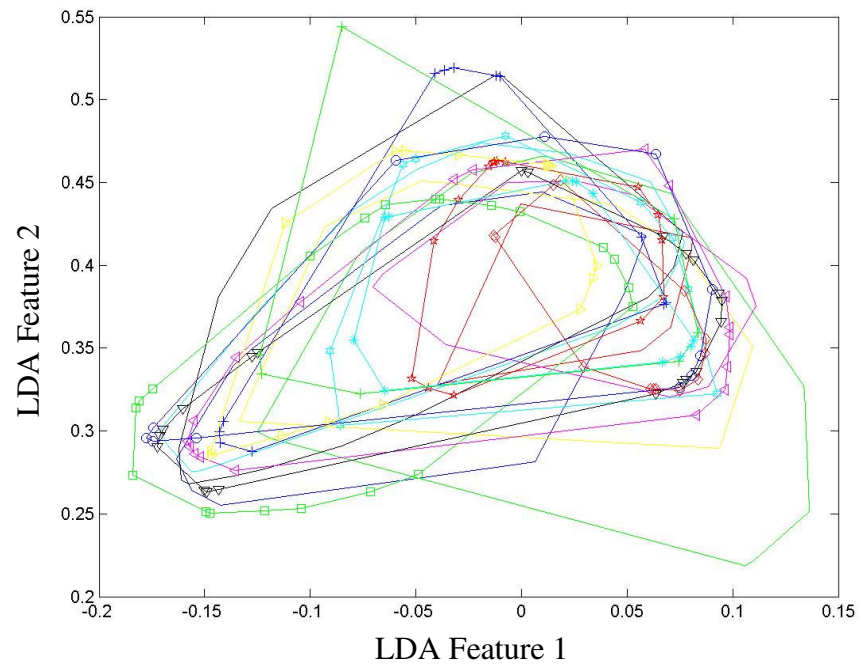
6.5 Data Visualization

6.5.1 Scatter Plot, Convex Hull and d^2 -display

In order to visualize the data, exploratory data tools were employed. An LDA transformation was applied to set 1 (Table 6.10) to yield a two-dimensional data set. The same LDA extractor was applied on set 2 (Table 6.10) to achieve a similar transformation. The features were used to create a scatter plot. The resulting scatter plots displayed the tremendous confusion between the various classes in both sets of data. In order to make the visualization easier, convex hulls were plotted for the data of each class. All these plots are displayed in Figure 6.4 and 6.5.

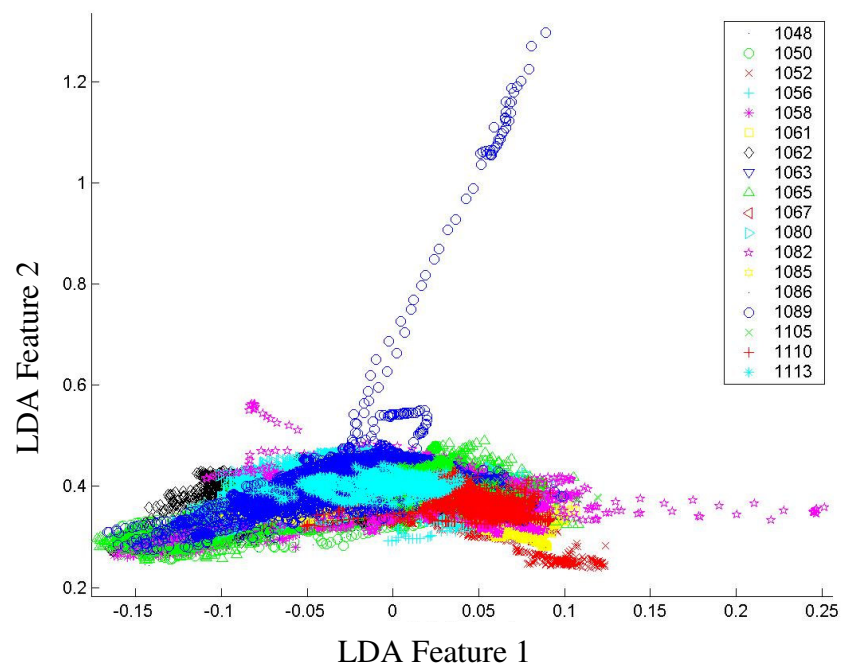


(a) Scatter plot

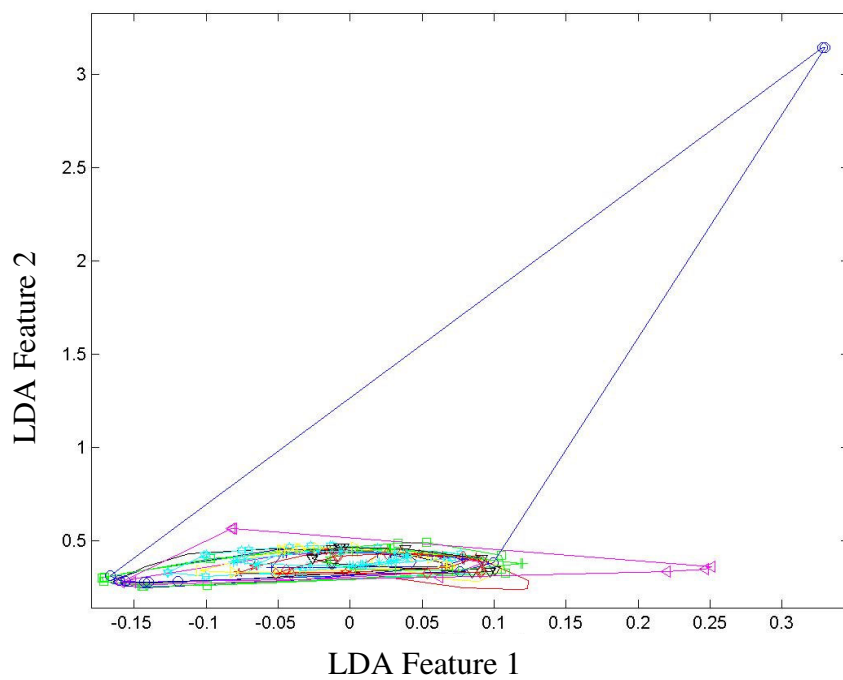


(b) Convex hull

Figure 6.4: Scatter plot and convex hull for set 1



(a) Scatter plot



(b) Convex hull

Figure 6.5: Scatter plot and convex hull for set 2

It was obvious from the convex hull plots that there was extensive overlap between the different classes resulting in the poor classification accuracy. The confusion reflected the complexity of the problem at hand.

The d^2 -display (Section 5.4.2) was used to determine the overlap between the largest class in the dataset, namely 1050, and the rest of the data. Set 1 from Table 6.10 was used for the purpose of this display. Figure 6.6 shows the large overlap. Further investigation into the dynamic range of values covered by class 1050 revealed that, it was effectively omni-present in the feature space. The cause was determined to be the coarse nature of the digital geological map. A closer look at the confusion matrix for error estimation of set 2 (Table 6.10) revealed that every class in the dataset was at least partially if not completely being confused with class 1050, hence affecting accuracy.

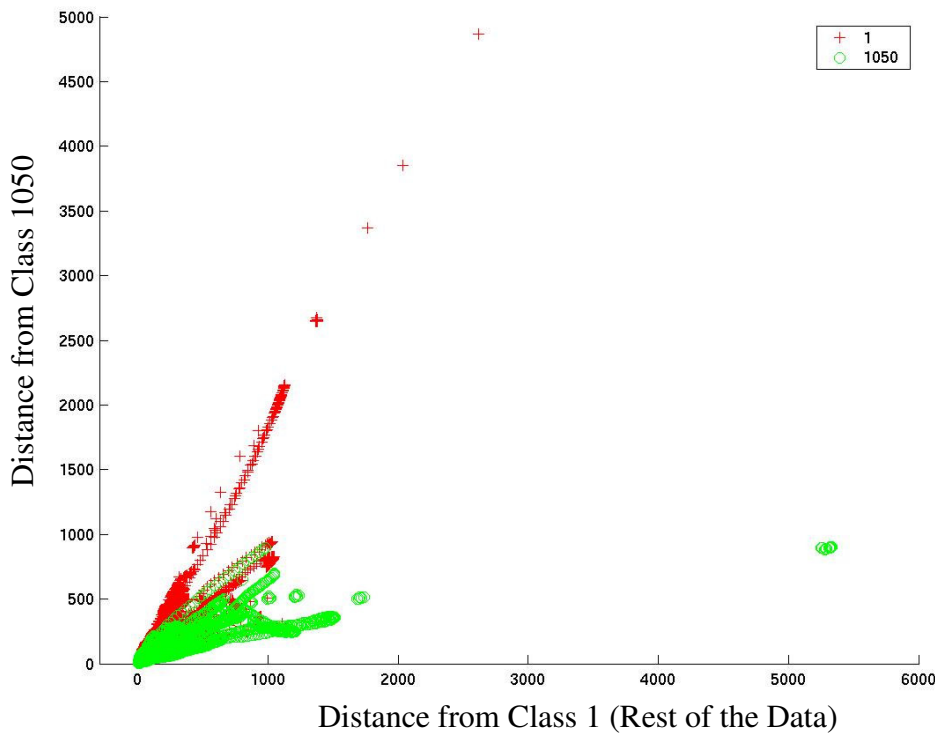


Figure 6.6: d^2 -display for class 1050 vs. rest of the data

6.5.2 Experiments without Class 1050

After visualizing the data, a decision was made to drop class 1050 from the dataset and proceed with the analysis. The datasets and results are in Table 6.12 and 6.13.

Lithological Unit	Set 1		Set 2		Set 3	
	No.	(%)	No.	(%)	No.	(%)
1048	4035	4.22	3999	4.17	3777	3.98
1052	1487	1.55	1472	1.54	1560	1.64
1056	10993	11.48	11112	11.60	11240	11.84
1058	8228	8.60	8055	8.41	8251	8.69
1061	10363	10.83	10466	10.92	10585	11.15
1062	5365	5.61	5845	6.10	5959	6.27
1063	1755	1.83	1694	1.77	1724	1.82
1065	9566	9.99	9245	9.65	9158	9.64
1067	2298	2.40	2614	2.73	2151	2.26
1080	5578	5.83	5649	5.90	5551	5.84
1082	14314	14.95	14481	15.12	13832	14.56
1085	3225	3.37	2999	3.13	3251	3.42
1086	2472	2.58	2146	2.24	2055	2.16
1089	9474	9.90	9857	10.29	9524	10.03
1105	2335	2.44	2381	2.49	2457	2.59
1110	2924	3.05	2744	2.86	2879	3.03
1113	1306	1.36	1046	1.09	1017	1.07

Table 6.12: Summary of lithological units for 3 ways split without class 1050

Feature Selection/ Classifier	Accuracy (%)	Improvement Over A <i>Priori</i> Probability (%)
SFS/Quad. Gaussian	21.38	6.82
SFS/1-NN	18.47	3.91
LDA/Quad. Gaussian	21.88	7.32
LDA/1-NN	18.92	4.40

Table 6.13: Classification accuracy for set 3 without 1050

As a result of the drop, a 7% increase was seen in the classification accuracy, over *a priori* probability.

6.6 Grouping of Units

6.6.1 Groups

Encouraged by the increase in the classification accuracy numbers, a grouping exercise was undertaken. We used geologic domain knowledge to find dissimilar units with similar magnetic signatures (e.g. units marked differently on the geologic map due to difference in age but possessing identical magnetic signatures). The confusion matrix derived during the process of estimating the error on set 2 (Table 6.12) was also consulted. Based on this matrix and on sets that made geologic sense, 5 major groups were derived. These groups are detailed in Table 6.14. The names of the lithological units are listed in Appendix B.

Groups	Lithological Units
1	1048, 1062, 1065
2	1058, 1063, 1067
3	1082, 1085, 1089, 1105, 1113
4	1056, 1080, 1086
5	1052, 1061, 1110

Table 6.14: Grouping of units

6.6.2 Experiments Conducted Post-Grouping

Once the grouping was completed, an experiment was conducted to determine the effect of the grouping on the accuracy numbers. Table 6.15 shows the data split between

the groups. As seen in Table 6.16, the grouping led to a significant increase in accuracy. Post grouping, the Gaussian classifier was able to achieve a 15.5% improvement over *a priori* probability.

Groups	Set 1		Set 2		Set 3	
	No.	(%)	No.	(%)	No.	(%)
1	18966	19.81	18894	19.89	19089	19.92
2	12281	12.83	12126	12.77	12363	12.90
3	30654	32.03	30081	31.67	30764	32.11
4	19043	19.89	18846	19.84	18907	19.73
5	14774	15.43	15024	15.82	14682	15.32

Table 6.15: Summary of lithological units after grouping

Feature Selection/ Classifier	Accuracy (%)	Improvement Over A <i>Priori</i> Probability (%)
SFS/Quad. Gaussian	47.57	15.46
SFS/1-NN	35.72	3.61
LDA/Quad. Gaussian	42.76	10.65
LDA/1-NN	32.52	0.41

Table 6.16: Classification accuracy after grouping

6.6.3 Experiments Using Pure Training Samples

To improve the classification accuracy further, another experiment was devised. Samples were included in the training set, set 1 (Table 6.15), only if all the points that went into the computation of the feature vector, belonged to the same class. In a sense, the training set now contained only “pure” samples. The experiment was conducted by keeping samples in set 1 (Table 6.15) pure and estimating the accuracy on set 3 (Table 6.15). As seen in Table 6.17, by curtailing the participation of impure samples in the

training set, the classification accuracy was increased to approximately 18% over *a priori* probability.

Feature Selection/ Classifier	Accuracy (%)	Improvement Over A <i>Priori</i> Probability (%)
SFS/Quad. Gaussian	50.03	17.92
SFS/1-NN	37.72	5.61
LDA/Quad. Gaussian	43.24	11.13
LDA/1-NN	32.52	0.86

Table 6.17: Classification accuracy for pure training samples

6.7 Effect of Varying Window Size

To study the effect of the window size on the classification accuracy, four window sizes were compared. Sets 1 and 2 (Table 6.15) were computed using increasing window sizes. The feature subset chosen in section 6.6.2 was used to conduct the experiments. The results are recorded in Table 6.18.

Window Size (m)	Accuracy (%)	Improvement Over A <i>Priori</i> Probability (%)
500	47.57	15.46
1000	47.28	15.17
1500	46.83	14.72
2000	46.94	14.83

Table 6.18: Effect of changing window size

There was no dramatic increase or decrease in the classification accuracy numbers due to change in window size. This could be attributed to the fact that as window size increases, the neighborhood in which the texture is computed increases, but there is also an increase

in the confusion due to the inclusion of several neighboring classes. There was no strong evidence to choose one window size over the other. The conclusion was that the heuristic rule, of not having a window size smaller than the smallest geological structure of interest, could be safely employed.

6.8 Effect of Varying Quantization Levels

More than half the texture measures computed were gray level co-occurrence matrix (GLCM) based measures. Hence, we studied the effect of the number of quantization levels employed to compute the GLCM-based measures, on the classification accuracy. Intuitively, it is known that coarse quantization using fewer gray levels results in faster computation of the measures and stable co-occurrence matrices, but results in a poor representation of the fine texture in the data. In contrast, fine quantization using more gray levels introduces longer computation times for the measures. It could also result in sparse and hence unstable co-occurrence matrices, rendering the computed measures useless to the classification problem. For the data used in this study, 256 quantization levels were found to be appropriate (Table 6.19). However, no generalization can be made regarding number of quantization levels that should be used. It is a data driven quantity that needs to be determined empirically.

Quantization levels	Accuracy (%)	Improvement Over A <i>Priori</i> Probability (%)
64	40.53	8.42
128	45.96	13.85
256	47.57	15.46
512	46.94	14.83

Table 6.19: Effect of the number of quantization levels

6.9 Comparative Performance of Features

No one feature or type of feature outperformed the others. Every type used in this study provided a few features that repeatedly showed up on the list of features chosen during the feature selection phase. We maintained a list of the features chosen each time during feature selection. This list is available in Appendix D. The set of features that occurred on more than 50% of the (7 out of 14) lists is provided in Table 6.20.

Feature	Occurrence Percentage (%)
First-order statistic: Mean	71
First-order statistic: Skewness	64
First-order statistic: Kurtosis	50
Second-order statistic: Correlation	71
Second-order statistic: Sum Average	50
Second-order statistic: Sum Entropy	57
Second-order statistic: Entropy	50
Second-order statistic: Difference Entropy	64
Gettings measure: No. of peaks and troughs	57
Gettings measure: Ratio of Euclidean length to No. of peaks and troughs	71
Texture spectrum measure: Black-white symmetry	50
Texture spectrum measure: Geometric symmetry	57
Fractal measure: Ratio of range to standard deviation (two-dimensional window)	64
Fractal measure: Hurst coefficient	50

Table 6.20: Best performing features

From the first-order statistics category, mean, skewness and kurtosis performed the best. Due to the dipolar nature of the magnetic anomaly, ideally, the mean should be zero for all lithological units and hence it should not be one of the measures chosen. But in order for the mean to approach zero, a window of infinite size is required. Practically speaking, the window sizes are limited. As a consequence, the mean value varies for different lithological units and proved to be a useful measure in this study. Skewness and kurtosis quantitatively capture the commonality in the shape of the magnetic anomalies. Since the shapes of magnetic anomalies for different lithological units vary, these features proved to be useful. In contrast, the other first-order statistics such as standard deviation and variance did not perform well. Variances of the different units do not vary much from each other in orders of magnitude.

From the second-order statistics category, correlation, sum average, sum entropy and entropy performed well. Entropy captures the disorder in the data. Since there is sufficient disorder or roughness in the aeromagnetic anomaly data, entropy and its variants proved to be useful measures. As to the failure of the other second-order statistics, the reasons can range from numerical methods used in their computation to the integrative effect (Figure 3.5) brought on by the summations that are present in almost every one of the measures.

In the Gettings measures category, number of peaks and troughs (NPT) and the ratio of Euclidean length to number of peaks and troughs (EL/NPT) performed better as compared to Euclidean length (EL). The reasons for this are illustrated in Figure 6.7. NPT is able to capture the undulating (rough) nature of the aeromagnetic data, while EL/NPT captures the undulating nature as well as the amplitude of the intensity values.

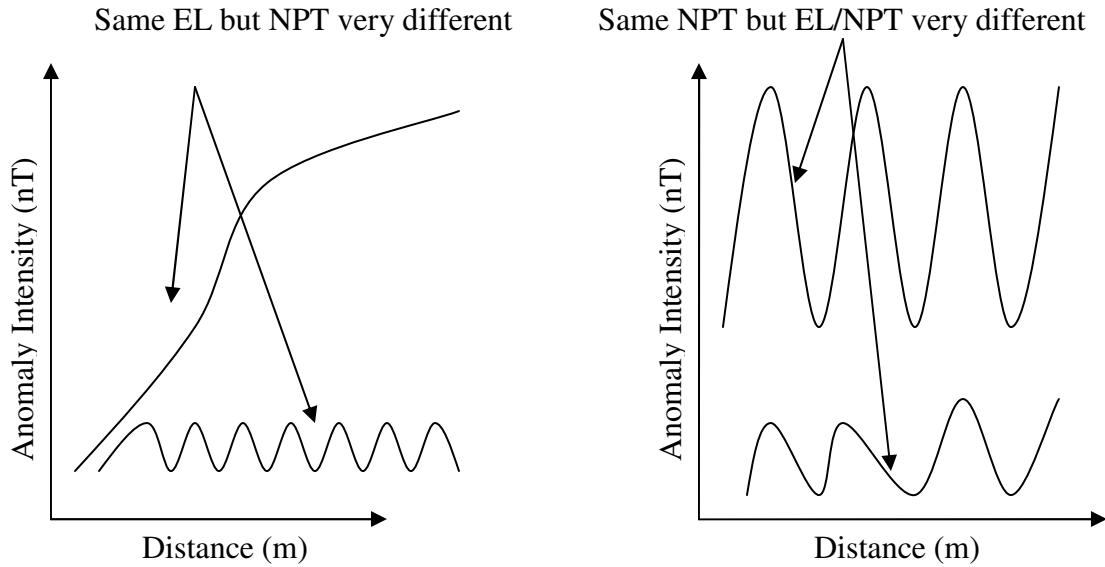


Figure 6.7: Reasons why EL and EL/NPT perform better than EL alone

Amongst the fractal dimension based features, the ratio of range to standard deviation and Hurst coefficient figured prominently in the list of best features, reinforcing the better performance of shape-based measures.

From the texture spectrum category, the measures attempting to capture macro texture, namely black-white symmetry and geometric symmetry, performed better than the ones attempting to capture micro-textures. In some sense these measures are two-dimensional versions of Getting's number of peaks and troughs measure and are successful for the same reasons.

6.10 Discussion of Results

6.10.1 Experimental Setup

We discovered that the experiments should be set up carefully in order to avoid bias in the results. Ideally, three spatially uncorrelated sets should be available for training, evaluation and testing. The aeromagnetic anomaly intensity values should be interpolated based on the sampling intervals targeted in the survey. The digital geological map of the region should be as detailed as possible to provide the best class information for the data. The window size used in computing the texture measures should not be smaller than the smallest geological structure of interest. It should also not be too large as that would result in multiple classes being included in the computation of the features. The quantization levels used in the gray-level co-occurrence matrix based features should be chosen such that the trade-off between quantization resolution and avoiding sparse co-occurrence matrices is balanced. For the survey used in this study, a window size of 500m and 256 quantization levels provided the best results.

6.10.2 Classification Accuracy

An 18% improvement over *a priori* probability was achieved in predicting lithological units for the Patagonia region. This statistically significant improvement over random guessing was achieved by grouping classes after analyzing the data at hand. The need to group classes resulted from the fact that the geological map, though being a state-of-the-art map for the region (map scale 1:125000), was coarse as compared to the aeromagnetic information available. Multiple lithological units in the region, uniquely labeled by the field geologist due to their varying age or chemical composition, have similar magnetic anomaly signatures. Hence, a machine vision system, without this

additional knowledge cannot successfully discriminate between the ungrouped classes. The moderate success of the classification algorithms, achieved after the grouping exercise, is a testimonial to the difficulty level of the problem as a whole. Due to the generality of the geologic data, as compared to the magnetic data, and overlap in the magnetic signature of several units, we can conclude that the accuracy results we obtained are a lower bound on the success of classification.

Overall, the quadratic Gaussian classifier performed better than the 1-NN classifier. A potential reason for this could be the severe confusion between the classes resulting in several incorrect minimum distance calculations. However, further investigation is required for a more definitive reason. The LDA transformation did not result in a dramatic increase in classification accuracy. This was probably because a linear combination of the measures, for projection into a lower dimensional space, was not sufficient to increase class separability.

6.10.3 Comparison with Previous Work

A few measures from each texture type performed well quantitatively. No one type dominated in its ability to quantitatively represent texture in the aeromagnetic intensity data. Within each type, some measures stood out and the probable reasons were discussed in section 6.9. Denitith's interpretation that GLCM-based and fractal dimension based features would work well, was quantitatively justified. However, his observation that simple statistical measures would not perform well, was contradicted by the presence of first-order statistics in the list of best performing features.

Getting's confidence in the quantitative ability of his wavelet-based measures was justified by the presence of 2 of his features on the best performers list. His first attempt

at an automated approach to delineating areas with common lithology was comprehensively extended. His observations regarding the nature of magnetic anomalies generated by lithological unit 1050, i.e., its positive as well as reversely polarized nature were corroborated in this study. Unit 1050 was found to have a presence in the entire range of anomalies encountered in this study and had to be dropped in order to perform continue the analysis with the rest of the units.

6.10.4 Comparison with Similar Quantitative Studies

Due to the lack of prior work in the area, it is difficult to express the significance of 18% improvement over *a priori* probability. Similar quantitative studies, conducted in the Humboldt river basin of Nevada, that employed weights-of-evidence methods to make quantitative geological predictions have not been able to do any better than 5% [50]. When put in this context, the results obtained in this study portend a promising future for automated lithological classification using aeromagnetic.

Chapter 7

CONCLUSION AND FUTURE WORK

Traditionally, amplitude and texture content of aeromagnetic anomaly images have provided useful geologic information. We undertook a quantitative study to estimate the capability of texture measures in automatically predicting the lithology of a region. Data from the Patagonia region of southeastern Arizona was used to setup a classification problem. We were able to demonstrate a significant improvement in classification accuracy over *a priori* probabilities. Comparative performance data for five types of texture measures was also obtained. To the best of our knowledge, this quantitative study is the first of its kind and it offers several useful insights into the process of automated lithological classification. Although, building an automated system with perfect accuracy remains a lofty goal, the exercise could definitely aid geophysicists and geologists in their pursuit to better understand the geology of a region.

This work can be extended in several ways. Different classifiers (neural nets, support vector machines) and different feature selection techniques (sequential backward search, branch and bound) can be used to try and improve classification accuracy. Adding wavelet-based measures to the already existing set of texture measures or performing the analysis solely in the wavelet domain, seems to hold promise. Another approach could be to include extra layers of information in the form of the gravity data or other such geophysical data for the region. Quantitative measures that suitably represent data from the new layers can be computed and added to the list of existing measures. The entire

study could also be redone with a fuzzy membership approach by relaxing the strict membership rules that are fundamental to statistical pattern classification. The grouping of lithological units that are geologically distinct but magnetically similar requires a closer look. “Magnetic lithologies” could be defined, using both the magnetic data and the geologic map, before conducting the classification experiments.

APPENDIX A: Creating a Composite Aeromagnetic Image

A.1 Loading Aeromagnetic Survey Data into Oasis Montaj

The database provided by the surveyor is in a *.XYZ format. This file format is native to the Oasis montaj software environment. Steps to load the file into Oasis are as follows

1. Launch Oasis montaj.
2. Create a new **Workspace** by clicking on the **File** menu and choosing **New Workspace**.
3. On the **Data** menu, click **Import**, and then click **ASCII**.

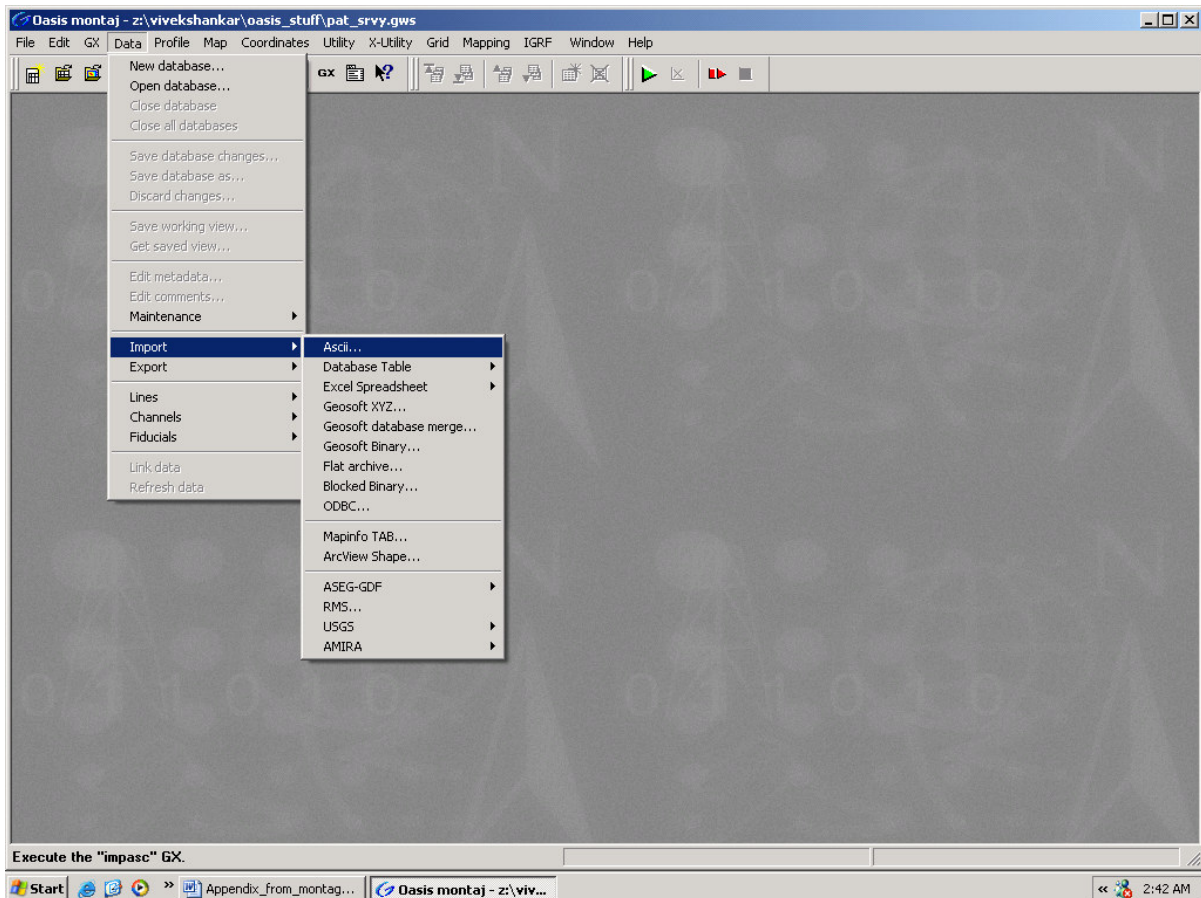


Figure: Screen shot of the Data menu.

4. Browse for the *.XYZ file provided by the surveyor and select it.
5. To navigate through the **Import Wizard**, click on **Wizard**.
6. Let the Wizard automatically find the row with headings and the delimiter used.
7. Next, choose the appropriate Column delimiter. Usually the default is correct.
8. Next, choose **Line** as the **Channel Type** for the first column, which is the Line column. For all other columns, let the **Channel Type** remain as **Data**.
9. Press **Finish**. The template need not be saved. Provide a database name. Set the **Maximum lines/group** to 500 and **Maximum channels/field** to 50. Click **OK**. Oasis will create a <database name>.gdb file. The file will be organized according to the survey lines. Note, due to the size of the file this process may take a while depending on the speed and memory available on the machine.

Fid	FLIGHT	date	time	GPSX	GPSY	latitude	longitude	gpsz
5037.0	6244.70	15	96322	22:44:41	531079	3466397	31.33325	-110.67331
5038.0	6244.80	15	96322	22:44:41	531088	3466396	31.33325	-110.67322
5039.0	6244.90	15	96322	22:44:41	531097	3466395	31.33324	-110.67312
5040.0	6245.00	15	96322	22:44:41	531106	3466394	31.33323	-110.67303
5041.0	6245.10	15	96322	22:44:41	531115	3466393	31.33322	-110.67294
5042.0	6245.20	15	96322	22:44:41	531124	3466392	31.33321	-110.67284
5043.0	6245.30	15	96322	22:44:41	531133	3466391	31.33320	-110.67274
5044.0	6245.40	15	96322	22:44:41	531142	3466390	31.33319	-110.67265
5045.0	6245.50	15	96322	22:44:41	531151	3466389	31.33318	-110.67256
5046.0	6245.60	15	96322	22:44:42	531160	3466388	31.33317	-110.67247
5047.0	6245.70	15	96322	22:44:42	531169	3466387	31.33316	-110.67237
5048.0	6245.80	15	96322	22:44:42	531178	3466386	31.33315	-110.67228
5049.0	6245.90	15	96322	22:44:42	531187	3466385	31.33314	-110.67218
5050.0	6246.00	15	96322	22:44:42	531195	3466384	31.33313	-110.67209

Summary Statistics: -50.6, 2529.0, 5108.6

Cell: 531259

For Help, press F1

Figure: Screen shot of the database file after it is loaded

A.2 Gridding Data Using Bi-Directional Cubic Spline

Oasis montaj offers several gridding techniques. Of these the bi-directional gridding scheme is best suited for aeromagnetic data. The gridding process is carried out in two steps. First, each line is interpolated along the original survey line to yield data values at the intersection of each required grid line with the observed line. The intersected points from each line are then interpolated in the between-line grid direction to produce a value at each required grid point. In both cases, along the flight line and across-line, the interpolation scheme can be independently selected. The various choices include linear, cubic (minimum curvature) or Akima spline. Splines are drafting aids used to draw smooth curves through a set of points. Many times a polynomial cannot be found which interpolates all the data points. Splines are piece-wise defined functions whose individual curves meet at the points. Cubic splines involve, fitting unique cubic polynomials between each of the data points, with the stipulation that the curve obtained must be continuous and appear smooth [45]. Steps to perform the bi-directional gridding are as follows

1. The <database name>.gdb file should be open.
2. On the **Coordinates** menu, click **Change X,Y coordinates**. Choose GPSX as **Current X** and GPSY as **Current Y** from the drop down options. Note, if the **Coordinates** menu is not visible it may be due to one of two reasons. Either the free-version of the software is running, in which case the license should be installed. Or, the **Load Custom Menu** icon on the toolbar should be clicked and the **Geophysics.omn** file should be chosen.

3. On the **Coordinates** menu, click **Set Projection**. Choose GPSX as the **X channel** and GPSY as the **Y channel**. Click on **Projection**. On the **Unknown Coordinate System** window, click **Modify**. Choose the **Projected(x,y)** as the **Coordinate System**. Click **Next**. The **Datum** and **Projection method** are chosen according to the specifications detailed in the survey document. For this study, the **Datum** was chosen as NAD27 and the **Projection method**, UTM zone 12N. The **local datum transform** was left as unknown. Choose the **length unit** as meter (or km depending on the units in which GPSX and GPSY are expressed). After confirming that all the inputs have been correctly accepted, click **OK** on the **Projected Coordinate System** window.
4. On the **Grid** menu, click on **Gridding**, followed by **Bi-directional line gridding** and **Dialog controls**. The **Channel to grid** is MAGIGRF (or the equivalent aeromagnetic anomaly channel). Provide a name for the **Output grid**. **Grid cell size** was chosen to be 10 m in this study. Click on the **Advanced** option. Choose, **Spline down-line** as Cubic and **across-line** as Cubic. Leave the rest of the options as default. Click on **Finish**.
5. <grid name>.grd can be viewed only as a map. To view the grid as a map, click on the **Grid** menu, followed by **Display grid** and **Single grid**. Browse to choose the <grid name>.grd. Leave the rest of the options as defaults. Click on **New Map**.

A.3 Generating a Database From a Grid

Once the dataset is uniformly re-sampled, it is available as a grid in the *.grd format, in the Oasis environment. But this format is not portable. Hence, the *.grd file needs to be

converted into a database or table that is available in a portable format, such as ASCII.

Steps to create a portable database are as follows

1. Click on the **Grid** menu, followed by **Utilities** and **Transpose**. Choose <grid name>.grd as the **Grid to transpose**. In the **Option**, choose force rows in X. This makes sure that the data is aligned such that points in every row line up.
2. Now, click on the **Grid** menu, followed by **Utilities** and **Save grid to database**. Choose <grid name>.grd as the **Grid File**. Provide a **New Database** name. Change the **Decimation Factor in the Y Direction** to 25. This ensures that in the y axis, the sampling interval is 250 m which is the average distance maintained during the survey. Leave the rest of the options as default.
3. Click on the **Data** menu, followed by **Export** and **Geosoft XYZ** options. Provide a XYZ file name. Click **OK**. Unselect the **Include Dummies** and **Include Line Header** options. Select the **CSV** option. Click **OK**. The database is now available in ASCII format to port to any application.

A.4 Loading a Coverage Into ArcGIS

The digitized version of the geological map of the region is available in the *.cov (or coverage) format [43]. This format is native to the ArcGIS software environment. A coverage is a format that uses a set of feature classes to represent geographic features. A feature class is an object that stores features and has a particular type of geometry. Each feature class stores a set of points, lines (arcs), polygons or annotation. Steps to load the coverage into ArcGIS as follows

1. Launch the **ArcMap** tool in the **ArcGIS** toolbox. Start with a new empty map.

2. On the **File** menu, click on **Add Data**. Choose the coverage file to be displayed. In this study the coverage file used was a subset of the i1109 coverage published by Drewes. It is called i1109_clip.cov.
3. Once the coverage is loaded, the appropriate symbol colors can be added for each of the lithological units. Right click on the file name and choose **Properties**. Under the **Symbology** tab, choose **Categories** followed by **Match to symbols in a style**. Choose the **Value field** as SYM_NUM. Browse to choose the style file **calcomp1.style**. This style file is part of the digital geological map published online. Click on **Add Values** and on **Complete List**. Select all the values and click **OK**. Now click on **Match Symbols**. Each of the lithological units should take on the designated symbol colors.

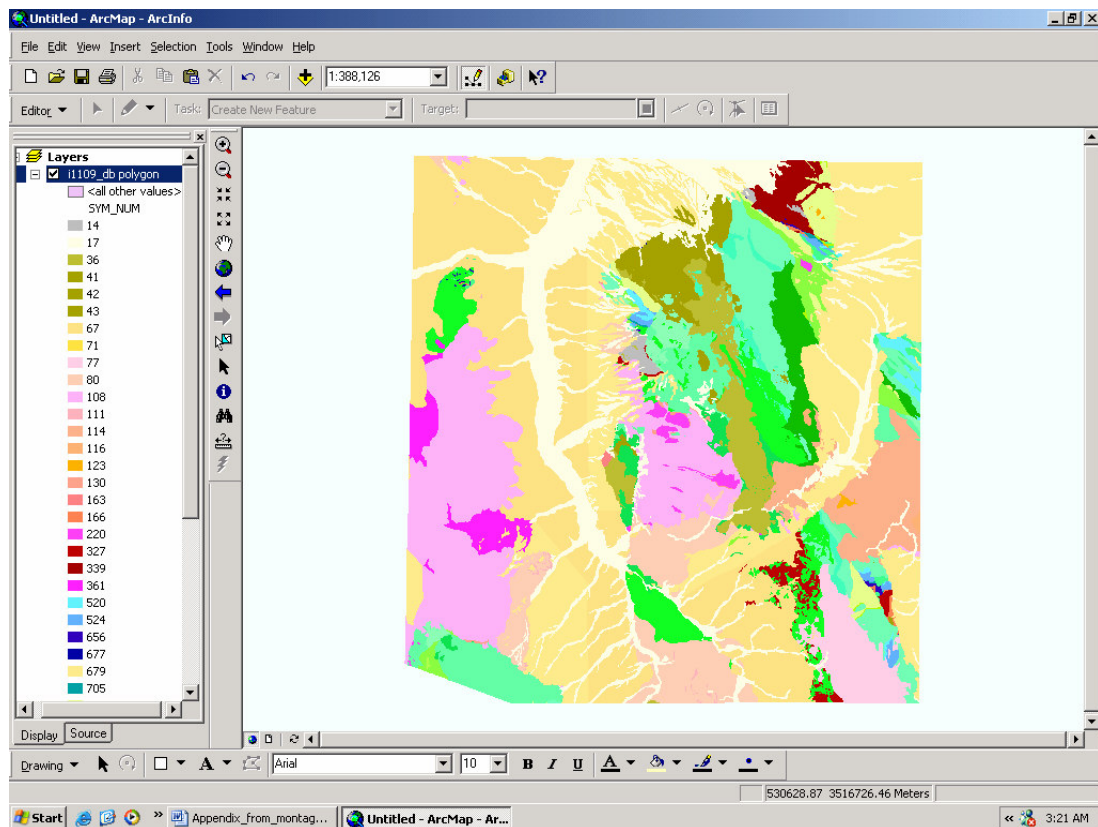


Figure: Screen shot of the ArcGIS environment after the coverage has been loaded.

A.5 Creating a Composite

Once the geological map of the region is loaded, the portable ASCII file containing the aeromagnetic anomaly information and the location information is imported into the ArcGIS environment. The two information sources are then intersected to create a single composite image, i.e., the *composite aeromagnetic image*. Steps to create the composite aeromagnetic image are as follows

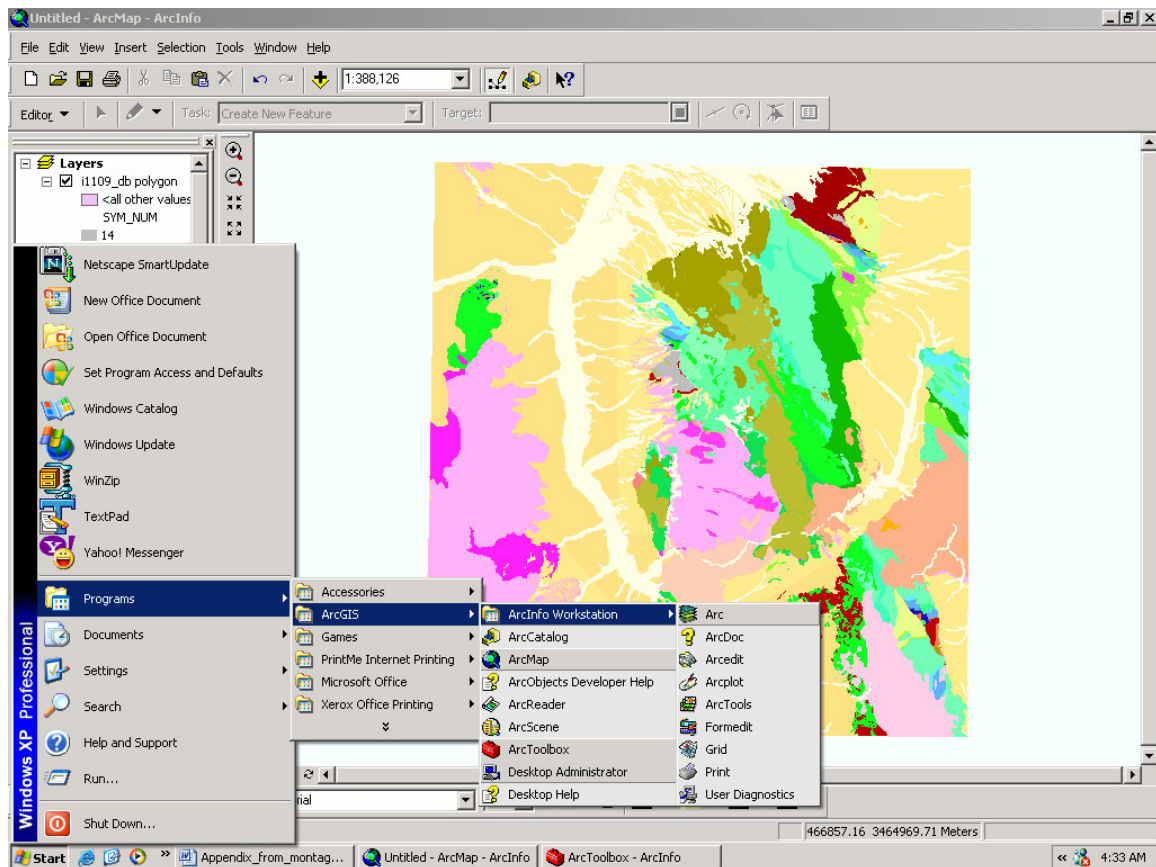


Figure: Screen shot showing all the components in the ArcGIS toolset

1. To import the aeromagnetic anomaly data into ArcGIS, click on the **Tools** menu, followed by **Add XY Data**. Browse and choose the ASCII file exported from Oasis. Note the exported oasis file must be given a *.asc extension in order to be recognized

by the ArcGIS environment. The **X field** should be GPSX and the **Y field**, GPSY. To provide a coordinate system, click on **Edit** and **Import**. Choose the i1109_clip file. Click on **Apply**. Finally click on **OK**. The aeromagnetic anomaly values are now loaded into ArcGIS in a temporary *.Events format.

2. From *.Events format, the data is first moved to an intermediate file format called a Shapefile. Right click on *.Events file and then on **Data**. Change the name from Export_Output.shp to some file name of less than 13 characters. Click **OK**.
3. From the ArcGIS Starup menu, launch the **ArcToolbox** component. Choose the **Shapefile to Coverage** option. Select intermediate shapefile as the Input Shapefile and provide a name for the coverage to be created, such as “aeromagcovfile.cov”.

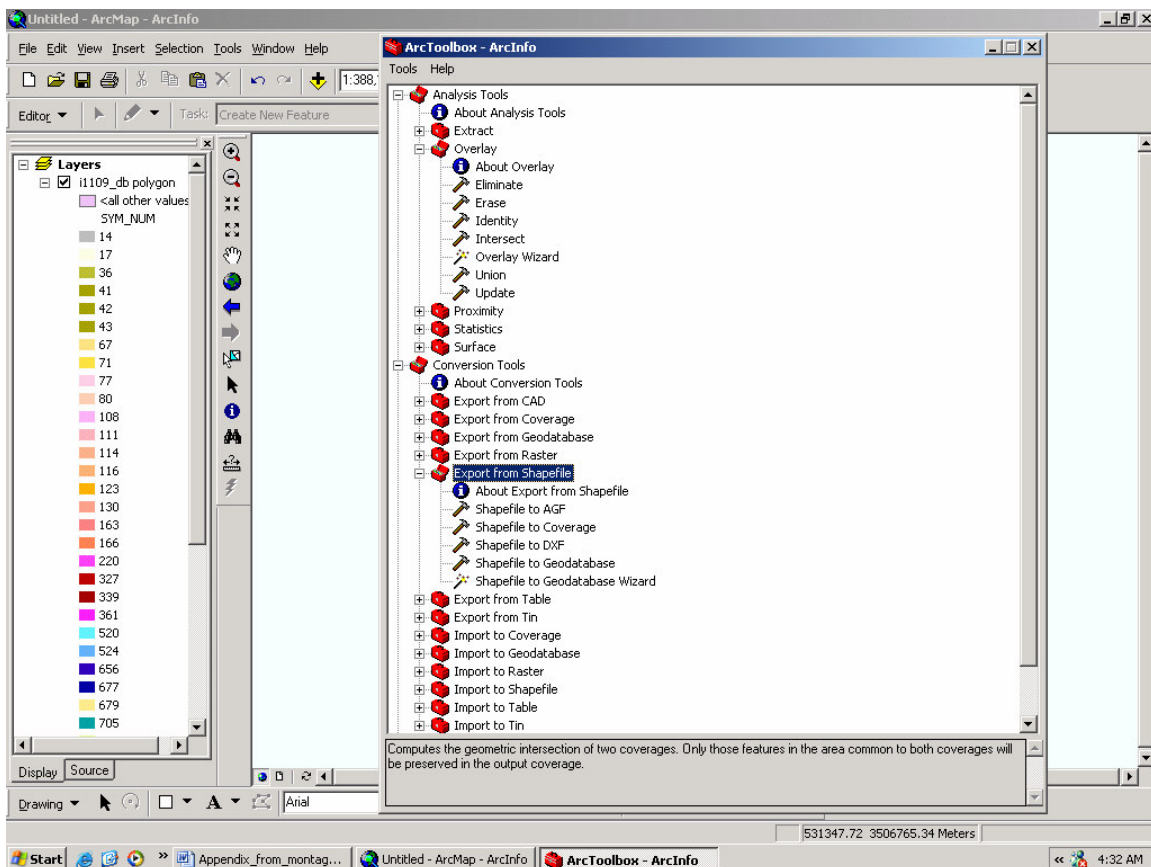


Figure: Screen shot of ArcToolbox

4. In the ArcToolbox, choose the **Intersect** tool from the **Overlay** set. Choose the <aeromagcovfile.cov> as the **Input coverage**. Select the i1109_clip as the **Intersect coverage**. Provide a name for the resulting coverage, such as “i1109aero.cov”. Once the intersect tool completes its work, launch the **ArcCatalog** component of the ArcGIS toolset. Browse and find <i1109aero.cov>. Right click and choose **Properties**. Select the **Projection** tab. Click on **Define**. Choose the option to **match existing data**. Click **Next**. Browse and find the i1109_clip coverage and choose it. Click **Next** and **Finish**. The composite aeromagnetic image is now ready.

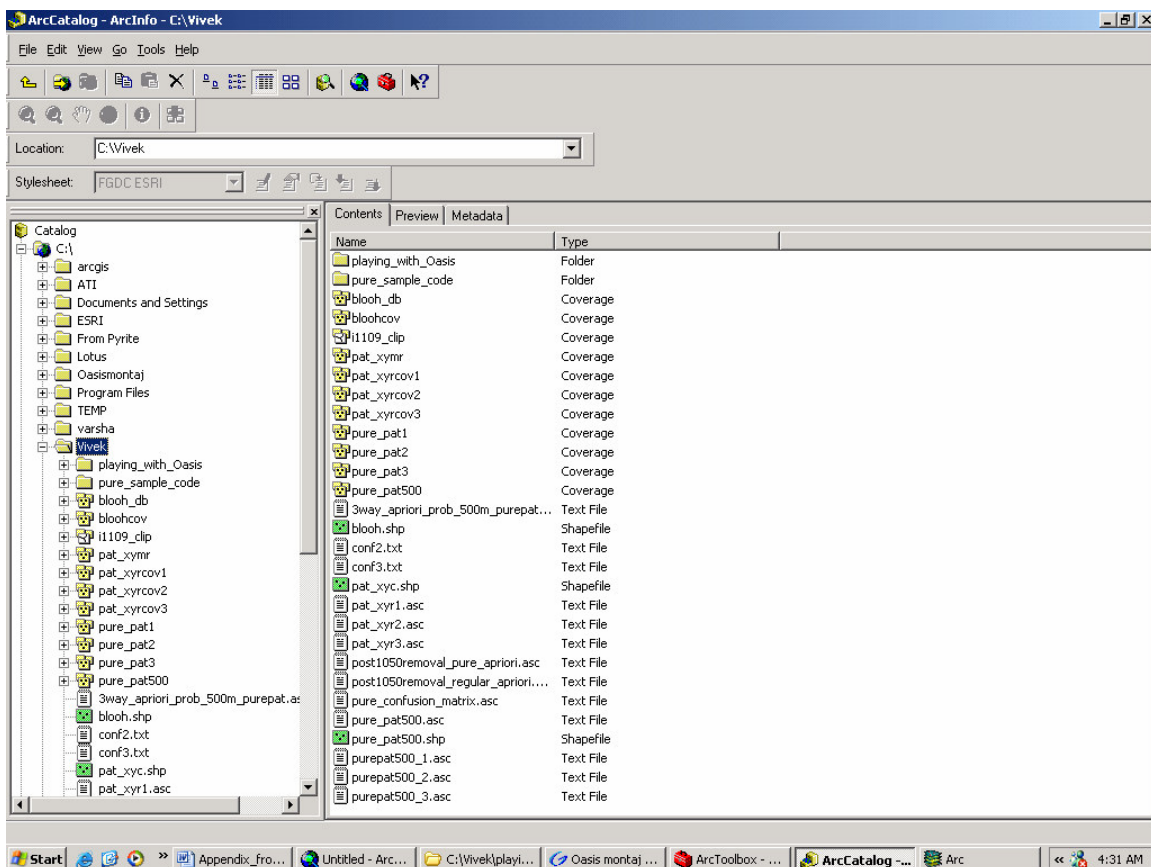


Figure: Screen shot of ArcCatalog

5. Finally, the composite aeromagnetic image has to be converted into a portable format such as ASCII. For that purpose, the **ArcInfoWorkStation** component of the ArcGIS

toolset is invoked. Choose the **Arc** option and a command line window appears with an Arc prompt. Several Unix like commands follow. These are listed.

- i. w <path for the current working directory>
- ii. Tables
- iii. sel <i1109aero.pat> (*.pat file of the intersected coverage)
- iv. items
- v. unload <outfile> <items> DELIMITED (the items include GPSX,GPSY,MAGIGRF and UNIT. UNIT is the unique number representing each class). Quit Tables and subsequently Arc by typing “q” at the prompt.

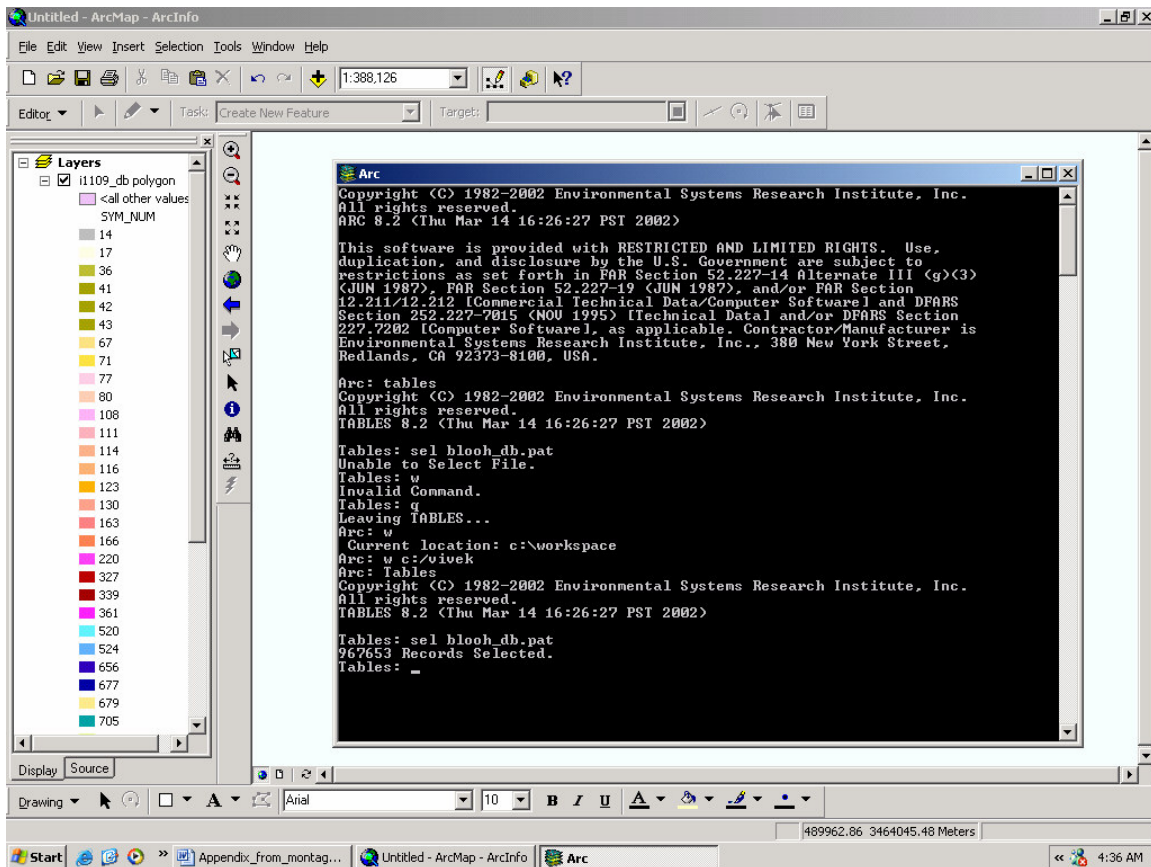


Figure: Screen shot of ArcInfoWorkStation

APPENDIX B: Lithological Units in Patagonia

B.1 Magnetic Rocks

Rocks with greater than 2000 samples (used in the study)

1. 1048 and 1049, “Tva”, “extrusive andesite and dacite (upper igneous and sedimentary rocks)”
2. 1050, “Tv”, “extrusive rhyolite and rhyodacite (upper igneous and sedimentary rocks)”
3. 1052, “Ti”, “intrusive rhyolite and rhyodacite (upper igneous and sedimentary rocks)”
4. 1056 and 1057, “Tlv”, “lower volcanic rocks (uppermost Cordilleran (Laramide) igneous rocks)”
5. 1058, “Tlg”, “lower granitoid rocks (uppermost Cordilleran (Laramide) igneous rocks)”
6. 1061, “Kd”, “diorite and quartz diorite (main Cordilleran (Laramide) igneous rocks)”
7. 1062, “Kq”, “quartz monzonite (main Cordilleran (Laramide) igneous rocks)”
8. 1063, “Kg”, “granodiorite (main Cordilleran (Laramide) igneous rocks)”
9. 1065 and 1066, “Kr”, “rhyodacite tuff and welded tuff (lower Cordilleran (Laramide) igneous and sedimentary rocks)”
10. 1067, “Ka”, “andesitic to dacitic volcanic breccia (lower Cordilleran (Laramide) igneous and sedimentary rocks)”
11. 1080 and 1081, “Klvs”, “andesitic to rhyolitic volcanic rocks, conglomerate, and sandstone (lower volcanic and sedimentary rocks)”

12. 1082, “Jg”, “stocks of pinkish-gray coarse-grained rock (granite and quartz monzonite)”
13. 1085, “JTRvs”, “rhyolitic tuff, welded tuff, lava, sandstone, and conglomerate (volcanic and sedimentary rocks)”
14. 1086, “TRm”, “stocks of dark-gray very coarse-grained monzonite and quartz monzonite (monzonitic rocks)”
15. 1089, “TRvs”, “rhyolitic to andesitic lava and pyroclastic rocks and intercalated sandstone, quartzite, and some conglomerate”
16. 1105 and 1106, “Yg”, “granodiorite and quartz monzonite (granitoid rocks)”
17. 1110, “Yc”, “Continental Granodiorite (granitoid rocks)”
18. 1113 and 1114, “Xp”, “Pinal Schist”

Rocks with fewer than 2000 samples (not included in the study)

1. 1045 and 1046, “Tb”, “basalt (upper igneous and sedimentary rocks)”
2. 1051, “Tug”, “granitoid rocks (upper igneous and sedimentary rocks)”
3. 1054, “Tg”, “granitoid rocks (uppermost Cordilleran (Laramide) igneous rocks)”
4. 1055, “Tlp”, “quartz latite porphyry (uppermost Cordilleran (Laramide) igneous rocks)”
5. 1060, “TKp”, “porphyritic and aplitic intrusive rocks (main Cordilleran (Laramide) igneous rocks)”
6. 1069, “Kuvs”, “volcanic and sedimentary rocks, undifferentiated (lower Cordilleran (Laramide) igneous and sedimentary rocks)”

B.2 Non-Magnetic Rocks

1. 1040, “Qg”, “gravel, sand, and silt (younger surficial deposits)”

2. 1041, “QTg”, “gravel, sand, and silt (older or undifferentiated surficial deposits)”
3. 1042, “QTgu”, “gravel, sand, and silt (older or undifferentiated surficial deposits)”
4. 1043, “Tuc”, “upper conglomerate, gravel, sand (older or undifferentiated surficial)”
5. 1053, “Tlc”, “lower conglomerate, gravel, and sand (upper igneous and sedimentary)”
6. 1064, “Kus”, “upper sedimentary rocks (lower Cordilleran (Laramide) igneous and sedimentary rocks)”
7. 1072, “Ks”, “sedimentary rocks (lowest Cordilleran (Laramide) sedimentary rocks)”
8. 1075 and 1076, “Kb”, “upper part of Bisbee Formation or Group, undifferentiated, and related rocks (Bisbee Formation or Group, undifferentiated)”
9. 1077, “Kbu”, “upper part of Bisbee Formation or Group, undifferentiated, and related rocks (Bisbee Formation or Group, undifferentiated)”
10. 1078, “Kbg”, “Glance Conglomerate of Bisbee Group or Glance Conglomerate of Bisbee”
11. 1087, “TRs”, “red mudstone, sandstone, and conglomerate, and intercalated rhyodacite volcanic rocks”
12. 1090, “Pzs”, “Rainvalley Formation to Bolsa Quartzite, undifferentiated”
13. 1094 and 1095, “Ps”, “sedimentary rocks of the Rainvalley Formation, Concha Limestone, and Scherrer Formation, undifferentiated (Naco Group)”
14. 1096, “PPs”, “sedimentary rocks of the Epitaph Dolomite, Colina Limestone, and Earp Formation, undifferentiated (Naco Group)”
15. 1098, “Ph”, “Horquilla Limestone (Naco Group)”
16. 1099, “MDs”, “Escabrosa Limestone and Martin Formation, undifferentiated”
17. 1100, “Cs”, “Abrigo Formation and Bolsa Quartzite, undifferentiated”

APPENDIX C: Feature Vector Elements

1. Aeromagnetic Intensity value	26. Entropy-m
First-Order Statistics	27. Entropy-r
2. Mean	
3. Variance	28. Difference Variance-m
4. Standard Deviation	29. Difference Variance-r
5. Coefficient of Variation	30. Difference Entropy-m
6. Skewness	31. Difference Entropy-r
7. Kurtosis	32. Information Measure of Correlation1-m
Fractal Dimension Based	33. Information Measure of Correlation1-r
8. Range/Standard Deviation (along flight line)	
9. Range/Standard Deviation (entire widow)	34. Information Measure of Correlation2-m
Second-Order Statistics (m-mean, r-range)	35. Information Measure of Correlation2-r
10. Angular Second Moment-m	
11. Angular Second Moment-r	36. Maximum Probability-m
12. Contrast-m	37. Maximum Probability-r
13. Contrast-r	Gettings Measures
	38. Euclidean Length
14. Correlation-m	39. Number of Peaks and Troughs
15. Correlation-r	40. Eucl. Length/No. of Peaks and Troughs
16. Sum of Squares-m	Texture Spectrum Measures
	41. Black-White Symmetry
17. Sum of Squares-r	42. Geometric Symmetry
18. Inverse Difference Moment-m	43. Degree of Direction
19. Inverse Difference Moment-r	44. Horizontal Symmetry
20. Sum Average-m	45. Vertical Symmetry
21. Sum Average-r	46. Diagonal Measure1 Symmetry
22. Sum Variance-m	47. Diagonal Measure2 Symmetry
23. Sum Variance-r	48. Central Symmetry
24. Sum Entropy-m	Fractal Dimension Based
	49. Hurst Coefficient
25. Sum-Entropy-r	

APPENDIX D: List of Features Chosen

Experiment	Feature List
Table 6.5 (SFS/ Gaussian)	2, 6, 7, 9, 14, 18, 24, 26, 32, 40, 44, 46
(SFS/ 1-NN)	1, 2, 6, 7, 14, 20, 24, 27, 28, 31, 36, 39, 40, 41, 42, 45, 49
Table 6.7 (SFS/Gaussian)	1, 6, 9, 10, 14, 18, 19, 26, 27, 31
(SFS/ 1-NN)	8, 7, 9, 18, 20, 24, 25, 28, 31, 39, 40, 41, 42
Table 6.9 (SFS/Gaussian)	2, 6, 14, 19, 31, 32, 36, 40, 42, 44, 46, 47
(SFS/ 1-NN)	1, 2, 6, 7, 9, 14, 25, 27, 39, 40, 41, 42, 49
Table 6.11 (SFS/Gaussian)	1, 2, 9, 24, 31, 33, 39, 47, 45, 49
(SFS/ 1-NN)	2, 6, 14, 19, 20, 24, 25, 27, 28, 40, 41, 42, 49
Table 6.13 (SFS/Gaussian)	1, 6, 7, 8, 9, 14, 18, 20, 26, 31, 32, 33, 37, 39, 44, 49
(SFS/ 1-NN)	2, 7, 8, 14, 20, 33, 36, 40, 41, 42, 44
Table 6.16 (SFS/Gaussian)	8, 9, 10, 11, 18, 20, 24, 31, 36, 37, 39, 40, 41, 43, 44, 49
(SFS/ 1-NN)	2, 6, 7, 9, 14, 25, 27, 31, 32, 36, 39, 40, 42, 45
Table 6.17 (SFS/Gaussian)	2, 6, 9, 24, 27, 31, 32, 42, 49
(SFS/ 1-NN)	1, 2, 14, 19, 20, 24, 25, 27, 31, 33, 36, 39, 40, 41, 44

REFERENCES

- [1] P. Kearey and M. Brooks, *An Introduction to Geophysical Exploration*, 2nd ed., Blackwell Scientific Publication Inc., Oxford, 1984.
- [2] M.C. Dentith et al., "Enhancement of subtle features in aeromagnetic data," *Exploration Geophysics*, vol. 31, pp. 104-106, March, 2000.
- [3] P.R. Milligan and P.J. Gunn, "Enhancement and presentation of airborne geophysical data," *Journal of Australian Geology and Geophysics*, vol. 17, no. 2, pp. 63-75, April, 1997.
- [4] M.C. Dentith, "Textural filtering of aeromagnetic data," *Exploration Geophysics*, vol. 26, pp. 209-214, September, 1995.
- [5] M.E. Gettings, "An interpretation of the 1996 aeromagnetic data for the Santa Cruz basin, Tumacacori mountains, Santa Rita mountains, and Patagonia mountains, south-central Arizona," *U.S. Geological Survey Open-File-Report 02-99*, (<http://geopubs.wr.usgs.gov/open-file/of02-99>), 2002.
- [6] D.P. Harris, *Mineral Resources Appraisal*, Oxford University Press, Oxford, 1984.
- [7] M.E. Gettings et al., "A complex systems model approach to quantified mineral resource appraisal," *Environmental Management*, vol. 33, no. 1, pp. 87-89, January, 2004.
- [8] P.S. Naidu and M.P. Mathew, *Analysis of Geophysical Potential Fields*, Elsevier Science B.V., Amsterdam, 1998.

- [9] F. Morrison et al., "The Berkeley course in applied geophysics," *Department of Civil and Environmental Engineering, University of California, Berkeley, California*, (<http://appliedgeophysics.berkeley.edu>), 2004.
- [10] C.R. Nave, "Hyperphysics," *Department of Physics and Astronomy, Georgia State University, Atlanta, Georgia*, (<http://hyperphysics.phy-astr.gsu.edu/hbase/emcon.html>), 2001.
- [11] D.P. Stern and M. Peredo, "The exploration of the Earth's magnetosphere," *NASA Goddard Space Flight Center, Greenbelt, Maryland*, (<http://www-spf.gsfc.nasa.gov/Education/Intro.html>), 2004.
- [12] R.J. Blakely et al., "Puget sound aeromagnetic maps and data," *U.S. Geological Survey Open-File-Report 99-514*, (<http://geopubs.wr.usgs.gov/open-file/of99-514>), 1999.
- [13] R.C. Brodie, "Airborne and ground magnetics," *Geophysical and Remote Sensing Methods for Regolith Exploration*, E. Papp (ed.), CRC LEME Open File Report 144, pp. 33-45, 2002.
- [14] A.K. Jain, *Fundamentals of Digital Image Processing*, 9th ed., Prentice-Hall, Englewood Cliffs, New Jersey, 1989.
- [15] R.C. Gonzalez and R.E. Woods, *Digital Image Processing*, 1st ed., Addison Wesley Inc., Reading, Massachusetts, 1992.
- [16] W.K. Pratt, *Digital Image Processing*, 2nd ed., John Wiley & Sons, New York, 1991.
- [17] R.M. Haralick et al., "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-3, no. 6, pp. 610-621, November, 1973.
- [18] S. Siggelkow, "Feature histograms for content-based image retrieval," *Computer Science Department, Institute for Pattern Recognition and Image Processing, Albert*

Ludwig University of Freiberg, Freiberg, Germany,

(<http://homepages.inf.ed.ac.uk/cgi/rbf/CVONLINE/entries.pl?TAG1096>), 2002.

[19] D. Csetverikov, "Basic algorithms for digital image analysis," *Institute of*

Informatics, Eotvos Lorand University, Budapest, Hungary,

(http://ssip2003.info.uvt.ro/lectures/chetverikov/shape_analysis.pdf), 2003.

[20] M. Tuceyran and A.K. Jain, "Texture analysis," *The Handbook of Pattern*

Recognition and Computer Vision, 2nd ed., C.H. Chen, L.F. Pau and P.S.P. Wang (eds.),

pp. 207-248, World Scientific Publishing Co., Hackensack, New Jersey, 1998.

[21] R. M. Haralick, "Statistical and structural approaches to texture," *Proceedings of the*

IEEE, vol. 67, no. 5, pp. 786-804, May, 1979.

[22] A.R. Rao, *A Taxonomy for Texture Description and Identification*, Springer-Verlag

Inc., New York, 1990.

[23] B. Julesz, "Experiments in the visual perception of texture," *Scientific American*,

vol. 232, pp. 34- 43, April, 1975.

[24] J.S. Weszka et al., "A comparative study of texture measures for terrain

classification," *IEEE Transactions on System, Man and Cybernetics*, vol. SMC-6, no. 4,

pp. 269-285, April, 1976.

[25] P.P. Ohanian and R.C. Dubes, "Performance evaluation for four classes of textural

features," *Pattern Recognition*, vol. 25, no. 8, pp. 819-833, August, 1992.

[26] A. Baraldi and F. Parmiggiani, "An investigation of the textural characteristics

associated with gray level cooccurrence matrix statistical parameters," *IEEE*

Transactions on Geoscience and Remote Sensing, vol. 33, no. 2, pp. 293-304, March,

1995.

- [27] D.C. He and L. Wang, "Texture unit, texture spectrum, and texture analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 4, pp. 509-512, July, 1990.
- [28] D.C. He and L. Wang, "Texture features based on texture spectrum," *Pattern Recognition*, vol. 24, no. 5, pp. 391-399, May, 1991.
- [29] A.P. Pentland, "Fractal-based description of natural scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, no. 6, November, 1984.
- [30] Y.Q. Chen and G. Bi, "On texture classification using fractal dimension," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 13, no. 6, pp. 929-943, September, 1999.
- [31] M. Pilkington, "Fractal magnetization of continental crust," *Geophysical Research Letters*, vol. 20, no. 7, pp. 627-630, April, 1993.
- [32] J.C. Russ, *Fractal Surfaces*, Plenum Press, New York, 1994.
- [33] R.O. Duda et al., *Pattern Classification*, 2nd ed., John Wiley & Sons, New York, 2002.
- [34] P.A. Devijver and J. Kittler, *Pattern Recognition: A Statistical Approach*, Prentice-Hall, London, 1982.
- [35] D.J. Hand, *Discrimination and Classification*, John Wiley & Sons, London, 1981.
- [36] K. Castleman et al., "Pattern recognition – Bayes risk pixel classifier," *Advanced Digital Imaging Research, LLC, League City, Texas*, (<http://adires.com/09/PatRec/PatRec.html>), 2002.
- [37] A.K. Jain et al., "Statistical pattern recognition: A review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4-37, January, 2000.

- [38] T.Y. Young and K.S. Fu (eds.), *Handbook of Pattern Recognition and Image Processing*, Academic Press, San Diego, 1986.
- [39] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed., Academic Press, San Diego, 1990.
- [40] A. Webb, *Statistical Pattern Recognition*, Oxford University Press, New York, 1999.
- [41] R.S. Sweeney, "Three aeromagnetic surveys in south central Arizona: A web site for distribution of data (on-line edition)," *U.S. Geological Survey Open-File-Report 00-155*, (<http://pubs.usgs.gov/of/2000/ofr-00-0155/>), 2000.
- [42] D.J. Dean, "The universal transverse Mercator coordinate system," *Geodesy, Cartography and Map Reading, College of Natural Resources, Colorado State University, Fort Collins, Colorado*, (http://www.cnr.colostate.edu/class_info/nr502/lg3/lg3_master.html), 2004.
- [43] H. Drewes, "Spatial Digital Database for the Tectonic Map of Southeast Arizona," *U.S. Geological Survey Miscellaneous Investigations Series Map I-1109*, (<http://pubs.usgs.gov/imap/i1109/>), 2002.
- [44] "Oasis montaj," *Geosoft Inc., Toronoto, Ontario, Canada*, (<http://www.geosoft.com/pinfo/oasismontaj/index.asp>).
- [45] S. McKinley and M. Levine, "Cubic spline interpolation," *Mathematics Department, College of the Redwoods, Eureka, California*, (<http://online.redwoods.cc.ca.us/instruct/darnold/laproj/Fall98/SkyMeg/Proj.pdf>), 1998.
- [46] "ArcGIS," *Environmental Systems Research Institute Inc., Redlands, California*, (<http://www.esri.com/software/arcgis/>).

- [47] “Image data for texture-based computer vision,” *Machine Vision Group, University of Oulu, Oulu, Finland*, (<http://www.ee.oulu.fi/research/imag/texture/texture.php>).
- [48] T.W. Rauber, “TOOLDIAG: Pattern recognition toolbox,” *Department of Informatics, Federal University of Espirito Santo, Vitoria, Brazil*, (<http://www.inf.ufes.br/~thomas/home/tooldiag.html>), 2001.
- [50] A.R. Wallace et al., “Assessment of metallic resources in the Humboldt river basin, northern Nevada,” *U.S. Geological Survey Bulletin 2218*, (<http://pubs.usgs.gov/bul/b2218>), 2004.